

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2003, 16:56:07 Search time 53 seconds

4306.581 Million cell updates/sec

Title: US-10-047-257-1

Percent score: 71

Sequences: 1 ATRHYLGVSLSDWQNSD.....NVRQALNEVLGCEQDLY 1438

Scoring table:

Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score as high as the one observed. The results are being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	7684	99.9	1457	20	B-domain deleted f
2	7681	100.0	1457	20	AAV21676
3	7684	99.9	1457	19	Human factor VIII
4	7684	99.9	1457	19	Human factor VIII
5	7684	99.9	1457	19	Human factor VIII
6	7674	99.8	1471	18	AAW3414
7	7659.5	99.8	1471	21	Human B-domain del
8	7659.5	99.6	1440	12	AAE12971
9	7656	99.2	1459	22	Factor VIII: C
10	7656	99.2	1459	22	Human factor VIII
11	7656	99.0	1459	22	AAE10833
12	7656	99.0	1459	22	Human factor VIII

Human factor VIII	10	7613	99.0	1459	22	AAE10827
Modified factor VI	11	7611.5	99.0	1516	9	AAE0265
Mutant mature factor VI	12	7592	98.7	1424	22	AAE04894
Human mature B-domain	13	7592	98.7	1424	22	AAE04894
Human mature B-domain	14	7592	98.7	1424	22	AAE04894
Modified factor VI	15	7581	98.6	1425	9	AAE0267
Modified factor VI	16	7578	98.5	1424	9	AAE0268
Human mature B-domain	17	7578	98.5	1445	23	AAE02540
Human mature B-domain	18	7578	98.5	1445	23	AAE02540
Human mature B-domain	19	7434	96.7	1661	18	AAE18670
Human mature B-domain	20	7353.5	95.6	1383	18	AAE13277
Procoagulant-activ	21	7345.5	95.5	1383	18	AAE13277
Procoagulant-activ	22	7345.5	95.5	1383	18	AAE13277
Procoagulant-activ	23	7345.5	95.5	1383	18	AAE13277
Human factor VIII	24	7234	94.1	2332	19	AAE05493
Human factor VIII	25	7234	94.1	2332	19	AAE05493
Human mature wild-	26	7234	94.1	2332	22	AAE0926
Human mature wild-	27	7234	94.1	2332	22	AAE0926
Human mature wild-	28	7234	94.1	2332	22	AAE0926
Human mature wild-	29	7234	94.1	2332	23	AAE0969
Human factor VIII	30	7234	94.1	2351	23	AAE21676
Factor VIII proc	31	7234	94.1	2351	23	AAE21676
Active factor VIII	32	7231	94.0	2351	18	AAE10527
Active factor VIII	33	7231	94.0	2351	18	AAE10527
Factor VIII: C (A9)	34	7230.5	94.0	2344	18	AAE1432
Active factor VIII	35	7230.5	94.0	2344	18	AAE1430
Active factor VIII	36	7230.5	94.0	2344	18	AAE1430
Human factor VIII	37	7230	94.0	2332	23	AAE0972
Human factor VIII	38	7230	94.0	2345	18	AAE1415
Active factor VIII	39	7230	94.0	2345	18	AAE1403
Active factor VIII	40	7229.5	94.0	2346	18	AAE1431
Active factor VIII	41	7229.5	94.0	2346	18	AAE1431
Active factor VIII	42	7229.5	94.0	2346	18	AAE1431
Active factor VIII	43	7229	94.0	2347	18	AAE1411
Active factor VIII	44	7229	94.0	2347	18	AAE1402
Factor VIII: C (A9)	45	7229	94.0	2351	18	AAE1532

ALIGNMENTS

RESULT 1
AAE01262
ID AAE01262 standard; protein; 1438 AA.

AC AAE01262;

XX 25-SEP-2000 (first entry)

DB B-domain deleted factor VIII sequence.

DT Factor VIII; procoagulant; adenovirus; adeno-associated strain;
RW Gene therapy; human Burkitt's lymphoma; HGB therapy;
KW Chimeric protein; Burkitt's lymphoma; Factor VIII; human;
XX

OS Homo sapiens.

XX

XX

XX

XX 08-DEC-1999; 99NOV023169.

XX 10-DEC-1998; 98US-0209516.

XX (PAB) BAYER CORP.

XX Cho M, Chan SY, Kelsey M, Yee H;

XX WPI; 2000-431311/37.

XX Producing cells expressing a protein having factor VIII procoagulant
XX activity especially, human factor VIII in an industrial scale, involve

XX 25-NOV-1998; 98MO-0825354.
 XX PF
 XX 16-JAN-1998; 98US-0071596.
 XX PR
 XX 05-DEC-1997; 97US-0067614.
 XX XX
 XX (IMMO-) IMMUNE RESPONSE CORP.
 XX XX
 XX Bialingmaier S, Gonzales JEN, Ili CR, Yang CQ;
 XX P1
 XX MPI; 1999-385602/32.
 XX DR
 XX M-7805; AAM82256, AAM82259, AAM82260.
 XX XX
 XX Genes and vectors exhibiting increased expression and novel splicing
 XX PT
 XX patterns, useful for expression of, e.g. beta-domain deleted factor
 XX FT
 XX VIII
 XX XX
 XX Disclousure; Page 72-78; 123pp; English.
 XX XX
 XX The invention describes novel genes and vectors exhibiting increased
 XX CC
 XX Factor VIII protein, that comprises one or more consensu
 XX CC
 XX consensus splice sites which have been corrected to increase expression.
 XX CC
 XX The method, DNA sequences and expression vectors can be used to increase
 XX CC
 XX modified 5' and/or 3' untranslated regions of a factor VIII gene. Genes containing
 XX CC
 XX levels and tissue-specific expression. The methods are used for
 XX CC
 XX identification and correction of consensus splice sites, addition of
 XX CC
 XX introns, optimization of 5' and 3' untranslated regions and increase in
 XX CC
 XX to treat a clinical disorder. to study RNA processing and/or gene
 XX CC
 XX regulation. The present sequence represents a beta-domain deleted factor
 XX CC
 XX VIII protein.
 XX XX
 XX Sequence 1457 AA;
 XX SQ

Query Match 100.0%; Score 7691; DB 20; Length 1457;
 Residual similarity 100.0%; Pred. MC. 0;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATATYTGAGLSMDGSLGELPVDRFPFRVPSFPRVSTVTKTLFVETVLEN 60
 QY 20 ATATYTGAGLSMDGSLGELPVDRFPFRVPSFPRVSTVTKTLFVETVLEN 79
 QY 61 TATKRPFMGLAGPTQAEVTVTVTLAKMAASHVPSGLVAVKASGAYDDTSQ 120
 QY 80 TAPKRPFMGLQPTQAEVTVTVTLAKMAASHVPSGLVAVKASGAYDDTSQ 139
 QY 121 REKEDAVPFGSHTVYMWLVKENGPMASDPLCTYSVLVDLVDLNSGLICALVR 180
 QY 140 REKEDAVPFGSHTVYMWLVKENGPMASDPLCTYSVLVDLVDLNSGLICALVR 199
 QY 181 EGSLAKCTLTHTLFLPAVDPGSKSHSTKSLMDRDAASNAMPQRTVNGVNR 240
 QY 200 EGSLAKCTLTHTLFLPAVDPGSKSHSTKSLMDRDAASNAMPQRTVNGVNR 259
 QY 241 SLUGLICRHSYTHVVGWGTTEVMSIFLEGHFTVNRQASLETSPITFATGLL 300
 QY 260 SLUGLICRHSYTHVVGWGTTEVMSIFLEGHFTVNRQASLETSPITFATGLL 319
 QY 301 NOLGQPLFLCHSSHQDGEHAYVKKVCSCEEPQLPMWNEAEVDDDTLSMDVVF 360
 QY 320 NOLGQPLFLCHSSHQDGEHAYVKKVCSCEEPQLPMWNEAEVDDDTLSMDVVF 379
 QY 361 DDNDSPEFTQLRSVAKKHPCVTHYKAAESRDHDAVLAADHRSVQVANKSPQIG 420
 QY 380 DDNDSPEFTQLRSVAKKHPCVTHYKAAESRDHDAVLAADHRSVQVANKSPQIG 439
 QY 421 KVKVKRNVATVETKRAHQSSGLGSLVGEVGTLLIFPKQASRTYVPHQI 480
 QY 440 KVKVKRNVATVETKRAHQSSGLGSLVGEVGTLLIFPKQASRTYVPHQI 497
 QY 481 TDVRLPLSHRLPAGVHLKDFPLFGEIFKFKVTVYEDGPTKSDPRCLTRYSSFVME 540

DB 500 TDVRLPLSHRLPAGVHLKDFPLFGEIFKFKVTVYEDGPTKSDPRCLTRYSSFVME 559
 QY 541 TDVRLPLSHRLPAGVHLKDFPLFGEIFKFKVTVYEDGPTKSDPRCLTRYSSFVME 600
 DB 560 KDLASGLICPLALCTKESVQKGNQNSDGNVLFLSPVDERESWLTENIQFLPNFAG 619
 QY 601 VOLDEPESANVNSINSTRVSLQSVCLSHLVNWTLSIGATDPLVFFSGTTFPH 660
 DB 620 VOLDEPESANVNSINSTRVSLQSVCLSHLVNWTLSIGATDPLVFFSGTTFPH 679
 QY 661 NVYEDTTLFFPFGSETVFNSENGVLWGLIAGNSDFRFGMTALLVSSCDKQTDYTE 720
 DB 680 NVYEDTTLFFPFGSETVFNSENGVLWGLIAGNSDFRFGMTALLVSSCDKQTDYTE 739
 QY 721 DSTVEDATLLSKNNAIPEPSGNPVLKGRQRTITLQSDQBEIDYDITSYEMKK 780
 DB 740 DSTVEDATLLSKNNAIPEPSGNPVLKGRQRTITLQSDQBEIDYDITSYEMKK 799
 QY 781 EDPDYVDENRNGSPFQKCTNFTVAVERLNDGNSSPHVLNRAQSGVPPKVV 840
 DB 800 EDPDYVDENRNGSPFQKCTNFTVAVERLNDGNSSPHVLNRAQSGVPPKVV 859
 QY 841 PQFTDQSTQPLVGELENHGLGLVTRAEVDNIVTFNQAQSPYSFSLSYEE 900
 DB 860 PQFTDQSTQPLVGELENHGLGLVTRAEVDNIVTFNQAQSPYSFSLSYEE 919
 QY 901 DQQAQAPRPNVPEKTKYTPKVOHNPVTDSPCKAWAYSDVLKDVHSLGLP 940
 DB 920 DQQAQAPRPNVPEKTKYTPKVOHNPVTDSPCKAWAYSDVLKDVHSLGLP 979
 QY 961 LVVCHTVLAPARGQTVQEPALFTTIDETKSYWTFENNERCAPCHIQEDPTFE 1020
 DB 980 LVVCHTVLAPARGQTVQEPALFTTIDETKSYWTFENNERCAPCHIQEDPTFE 1039
 QY 1021 NYEPHANXIMDTLGLVWQADQRIWVLLSWGSENIHSIHESGVTVTKKXYA 1080
 DB 1040 NYEPHANXIMDTLGLVWQADQRIWVLLSWGSENIHSIHESGVTVTKKXYA 1099
 QY 1081 LVYLVPGVFFTEVMSAGVWVCLIGVHLQNGSLFLVGNWCVTVGNAASHID 1140
 DB 1100 LVYLVPGVFFTEVMSAGVWVCLIGVHLQNGSLFLVGNWCVTVGNAASHID 1159
 QY 1141 VQTAGSQVQWAPLAVKSGSINANSTLEPSPKVLVLAHLHNGIKQQAQKES 1200
 DB 1160 VQTAGSQVQWAPLAVKSGSINANSTLEPSPKVLVLAHLHNGIKQQAQKES 1219
 QY 1201 ELYTVQIMVSLDQKQVTKNGSTQTLVTFNVDSSGIDKHNINFTPIARYLRHPT 1260
 DB 1220 ELYTVQIMVSLDQKQVTKNGSTQTLVTFNVDSSGIDKHNINFTPIARYLRHPT 1279
 QY 1261 ITRISIRLMEIACLNCSNPLQWESLQASQITASSITFMFASPSKARHLAQ 1320
 DB 1280 ITRISIRLMEIACLNCSNPLQWESLQASQITASSITFMFASPSKARHLAQ 1339
 QY 1321 ISNAPRNPQNNPHELOVQKMYTVGVTVQNSLISLNTVKEPLISSQOQGTFLF 1380
 DB 1340 ISNAPRNPQNNPHELOVQKMYTVGVTVQNSLISLNTVKEPLISSQOQGTFLF 1399
 QY 1381 PQNGVVPQNGQSTFPFVNSLPPCLTRLLHPSQNVWDIQLAMNVLGCEADLY 1438
 DB 1400 PQNGVVPQNGQSTFPFVNSLPPCLTRLLHPSQNVWDIQLAMNVLGCEADLY 1457

RESULT 3
 AAM46246
 ID AAM46246
 AA AAM46246 standard; Protein; 1457 AA.
 CC
 XX AAM46246;
 XX 25-MAR-2003 (updated)
 XX 06-AUG-1998 (first entry)
 XX

Human factor VIII beta-domain deleted SQN deletion protein sequence.

Replication defective; recombinant retrovirus; RV; therapeutic protein; haemophilia; thrombosis; hypercoagulable disorder; liver disease; human; hepatitis; chalasemia; phenylketonuria; Leech-Nyhan syndrome; diabetes; hypophosphatemia; adenosine deaminase deficiency; HIV infection; anemia; Guacher's syndrome; high blood pressure; Alzheimer's disease; autoimmune; inflammatory disease; factor VIII.

OS Homo sapiens.

XX MO800541-A2.

XX 08-JUN-1998.

XX 02-JUL-1997; 97MO-US117784.

XX 04-JUN-1987; 97US-0869309.

XX 03-JUL-1996; 96US-0645601.

XX 13-AUG-1996; 96US-0696381.

XX (CHIR) CHIRON CORP.

XX Jolly DJ, Barber JR, Chang SMW, Respass JG, Allen JR, Boder M; Chong K, De La Vega D, Depolo JR, Hau DC, Ibanes CE; Mitcattedel DM, Prussack CE, Greengard J, Lee R; WPI: 1998-08696/08. N-PS08; MAV19581.

XX New replication defective recombinant retro-viruses - which can be administered to provide long term systemic expression of therapeutic protein in blood, useful in, e.g. treating hyper-coagulable disorders

XX Example 28; Pages 213-217; 272pp; English.

XX This is the beta-domain deleted SQN deletion protein of human factor VIII. The beta-domain deleted SQN deletion protein is a retroviral vector expressing human factor VIII. The invention provides a preparation of replication defective recombinant retrovirus (RV) expressing a therapeutic protein. The RV preparation is resistant to degradation by human complement and is capable of inducing long term expression of the therapeutic protein in a human. The long term measurable level of the therapeutic protein being produced in the blood of the human for a period of at least 30 days after the administration of the recombinant retrovirus. The long term measurable level of the therapeutic protein to treat, e.g. haemophilia A, haemophilia B, thrombosis, hypercoagulable disorders, liver diseases such as hepatitis, disorders such as thalassemia, phenylketonuria, Leech-Nyhan syndrome, muscular dystrophy, inherited emphysema, familial hypercholesterolemia, diabetes, hypothyroidism, adenine deaminase deficiency, alphas-1 antitrypsin deficiency, Guacher's syndrome, anemia. Infections such as influenza, high blood pressure, autoimmune disease, autoimmune of infection, and cancer. The invention also provides a method of surviving inactivation in human serum thereby allowing efficient gene transfer over prolonged periods of time.

XX Updated on 25-MAR-2003 to correct FI field.)

XX Sequence 1457 AA;

Query Match 99.98; Score 7694; DB 19; Length 1457; Similarity 99.98; Percent Identical 99.98; Mismatches 1; Indels 0; Gaps 0; Matches 1437; Conservative

QY 1 ATNRYIGLAVELNDVMDGLDGLVDARPPVPKSPFPEVYVTKTLFVEFTVILFN 60

2 ATNRYIGLAVELNDVMDGLDGLVDARPPVPKSPFPEVYVTKTLFVEFTVILFN 79

DB 61 TAKERPPWMLGLFTQAEYDVTVDITDGMASHFSLHVGVSFKASGEAGVDDQSQ 120

80 TAKERPPWMLGLFTQAEYDVTVDITDGMASHFSLHVGVSFKASGEAGVDDQSQ 139

121 REKEDNDVFGSHSTVYVQAKNGWPAADPCLCTSTLSDVDVLDNSGLIGALLVCR 180

140 REKEDNDVFGSHSTVYVQAKNGWPAADPCLCTSTLSDVDVLDNSGLIGALLVCR 199

181 REGLAKETQDLKFLFLFANVDEGKGSMSSTGSLADQSDAANAPQGTGVNVR 240

200 REGLAKETQDLKFLFLFANVDEGKGSMSSTGSLADQSDAANAPQGTGVNVR 259

241 SLGCLTCHRSYVTVAVTNGTFTVHSIFLEGHFLVDRHQASLEISPTFTTAQTL 300

260 SLGCLTCHRSYVTVAVTNGTFTVHSIFLEGHFLVDRHQASLEISPTFTTAQTL 319

301 MDGQGLFLCHSSHQHQENYVQVSCPEELPKQMKNEEADYDDDLTDSMDVRF 360

320 MDGQGLFLCHSSHQHQENYVQVSCPEELPKQMKNEEADYDDDLTDSMDVRF 379

361 DDNSPSFTQLRSVAKQHPKTVHYTAAEEDNDVAVLPDQBSYQVUNSGPQIG 420

380 DDNSPSFTQLRSVAKQHPKTVHYTAAEEDNDVAVLPDQBSYQVUNSGPQIG 439

421 RYKVEYVNAVTDTEKTRALQHSGLIGLQVSGVGLTLTFKQASRPVYVPHGI 480

440 RYKVEYVNAVTDTEKTRALQHSGLIGLQVSGVGLTLTFKQASRPVYVPHGI 499

481 TDYRSLYSLRPGVGLADPFLPGEIKPKVQVTVGQDQKQDQVYVSPFQNG 540

500 TDYRSLYSLRPGVGLADPFLPGEIKPKVQVTVGQDQKQDQVYVSPFQNG 559

541 PDYRSLYSLRPGVGLADPFLPGEIKPKVQVTVGQDQKQDQVYVSPFQNG 600

560 PDYRSLYSLRPGVGLADPFLPGEIKPKVQVTVGQDQKQDQVYVSPFQNG 619

601 VQLEPDSFQSNHNSINGVTVDSQLSLVCLHEHNAVYLLSGADDTPLVYVSGV 660

620 VQLEPDSFQSNHNSINGVTVDSQLSLVCLHEHNAVYLLSGADDTPLVYVSGV 679

661 NVYEDVTLFPFGSGTVMNSMPLWLCQNSDFNRMGTALLVSCDKNTDYTE 720

680 NVYEDVTLFPFGSGTVMNSMPLWLCQNSDFNRMGTALLVSCDKNTDYTE 739

721 DSYEDVTLFPFGSGTVMNSMPLWLCQNSDFNRMGTALLVSCDKNTDYTE 780

740 DSYEDVTLFPFGSGTVMNSMPLWLCQNSDFNRMGTALLVSCDKNTDYTE 799

781 DSYEDVTLFPFGSGTVMNSMPLWLCQNSDFNRMGTALLVSCDKNTDYTE 840

800 DSYEDVTLFPFGSGTVMNSMPLWLCQNSDFNRMGTALLVSCDKNTDYTE 859

841 DSYEDVTLFPFGSGTVMNSMPLWLCQNSDFNRMGTALLVSCDKNTDYTE 900

860 DSYEDVTLFPFGSGTVMNSMPLWLCQNSDFNRMGTALLVSCDKNTDYTE 919

901 DSYEDVTLFPFGSGTVMNSMPLWLCQNSDFNRMGTALLVSCDKNTDYTE 960

920 DSYEDVTLFPFGSGTVMNSMPLWLCQNSDFNRMGTALLVSCDKNTDYTE 979

961 DSYEDVTLFPFGSGTVMNSMPLWLCQNSDFNRMGTALLVSCDKNTDYTE 1020

980 DSYEDVTLFPFGSGTVMNSMPLWLCQNSDFNRMGTALLVSCDKNTDYTE 1039

1021 DSYEDVTLFPFGSGTVMNSMPLWLCQNSDFNRMGTALLVSCDKNTDYTE 1080

1040 DSYEDVTLFPFGSGTVMNSMPLWLCQNSDFNRMGTALLVSCDKNTDYTE 1099

1081 DSYEDVTLFPFGSGTVMNSMPLWLCQNSDFNRMGTALLVSCDKNTDYTE 1140

1100 DSYEDVTLFPFGSGTVMNSMPLWLCQNSDFNRMGTALLVSCDKNTDYTE 1159

1141 DSYEDVTLFPFGSGTVMNSMPLWLCQNSDFNRMGTALLVSCDKNTDYTE 1200

541 RUASGLTIGELLICWESQKNGKNGMKNWTLFQVDFWNSWTLFENKQVFNHNSG 600
 Db 560 RUASGLTIGELLICWESQKNGKNGMKNWTLFQVDFWNSWTLFENKQVFNHNSG 619
 Qy 601 VQLEDFEQANIKMSIKVDFDGLQALQWKNWTLFICGADQVDFVFFSWFPH 660
 Db 620 VQLEDFEQANIKMSIKVDFDGLQALQWKNWTLFICGADQVDFVFFSWFPH 679
 Qy 661 KVVYEDTLTLFFSSEVFNSMENPQJLIGCHNSDFNRCMTALVNSCDKNDYVE 720
 Db 680 KVVYEDTLTLFFSSEVFNSMENPQJLIGCHNSDFNRCMTALVNSCDKNDYVE 739
 Qy 721 IDIEDIGAVLLSQNNALFFSPFSQK-----DPVLKHQDRIETLQASQVE 766
 Db 740 IDIEDIGAVLLSQNNALFFSPFSQK-----DPVLKHQDRIETLQASQVE 799
 Qy 767 EYVDYDITLVEMKEDDITDDENQSPRSFQKTRHFIATVEKLMYDGNSSPFLVN 826
 Db 800 EYVDYDITLVEMKEDDITDDENQSPRSFQKTRHFIATVEKLMYDGNSSPFLVN 859
 Qy 827 RAQSQGVDFQVDFQVDFQVDFQVDFQVDFQVDFQVDFQVDFQVDFQVDFQ 886
 Db 860 RAQSQGVDFQVDFQVDFQVDFQVDFQVDFQVDFQVDFQVDFQVDFQ 919
 Qy 887 RYVFSYSLISYEDQKQKPRKQVFNKPKETKFNKVKOHMAVTFDFCKAMAFSD 946
 Db 920 RYVFSYSLISYEDQKQKPRKQVFNKPKETKFNKVKOHMAVTFDFCKAMAFSD 979
 Qy 947 VQLEKQVHSLIGELLICWNTINPAPKQVDFQVDFQVDFQVDFQVDFQVDF 1006
 Db 980 VQLEKQVHSLIGELLICWNTINPAPKQVDFQVDFQVDFQVDFQVDFQVDF 1039
 Qy 1007 APCNQIMEDPTFKENVEFHAIQVIMDTLQGLWAQDQIRMTLLSGNSNHSIHFS 1066
 Db 1040 APCNQIMEDPTFKENVEFHAIQVIMDTLQGLWAQDQIRMTLLSGNSNHSIHFS 1099
 Qy 1067 HYPVVKKEETKVALMPLPVFETVEMLPSTAGVWVCLIGELHAGNSTLFTVSNK 1126
 Db 1100 HYPVVKKEETKVALMPLPVFETVEMLPSTAGVWVCLIGELHAGNSTLFTVSNK 1159
 Qy 1127 CQFLGWASGHIDPQVTRASQVQWAPLHFLVSGTNSMSTKPSNKVQLLAPMI 1186
 Db 1160 CQFLGWASGHIDPQVTRASQVQWAPLHFLVSGTNSMSTKPSNKVQLLAPMI 1219
 Qy 1187 HGTGQVQAKQKPSLSLISQVTLQGLQKQVQVQVQVQVQVQVQVQVQVQV 1246
 Db 1220 HGTGQVQAKQKPSLSLISQVTLQGLQKQVQVQVQVQVQVQVQVQVQVQV 1279
 Qy 1247 NPPTATVRLHHTVHTSTELPWEKACWASCSNPLGCHESLSDAQTSASVYTHNE 1306
 Db 1280 NPPTATVRLHHTVHTSTELPWEKACWASCSNPLGCHESLSDAQTSASVYTHNE 1339
 Qy 1307 RYNSKQVHLLHCHSNRQKQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 1366
 Db 1340 RYNSKQVHLLHCHSNRQKQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 1399
 Qy 1367 LIGSSQV 1426
 Db 1400 LIGSSQV 1459
 Qy 1427 MWVQGEAQDLY 1438
 Db 1459 MWVQGEAQDLY 1471

RESULT 7

ID AAKI2971

ID AAKI2971 standard; protein; 1440 AA.

XX AAKI2971;

AC 25-MAR-2003 (updated)

DT 09-JAN-2003 (updated)

02-OCT-1991 (first entry)
 Factor VIII:SQ.

Factor VIII; B domain; haemophilia.

Unidentified.

W03109122-A.

27-JUN-1991.

06-DEC-1990; 90WO-SB00809.

15-DEC-1989; 89SE-0004239.

(KAB1) KABIVITRUB AB.

Almeredt AB, Hellerum EM, Larsson K, Lind P, Sandberg HI;

Spira J, Sydowback PM;

MF1; 1991-208149/28.

Recombinant human factor VIII deriv, deoxyribonucleic acid -

encoding protein comprising two chains linked by segment of B domain.

Disclousure, Fig 1, 35pp, English.

The protein is a fusion between Phe 742 and Ser 1637 of the factor

VIII protein (factor VIII:SQ). In order to produce a

factor VIII deletion derivative that can be produced in vivo and/or

in vitro, a DNA construct was prepared containing a fusion of

90 kb and 80 kb the amino acid sequences surrounding Arg 740 and

Arg 1648 have to be conserved in order to preserve the structural

requirements for correct cleavage. In this example, amino acids 743

new polypeptide chain is obtained where there are 14 amino acids. A

linking Arg 740 and Arg 1648. Of these 14 amino acids, the sequence

of the five N-terminal ones directly corresponds to the five amino

acids following Arg 740 in full-length factor VIII. Also, the sequence

directly corresponding to the 12 amino acids preceding Glu 1689 in full-

length factor VIII, thus creating a 3 amino acid overlap between the

N- and C-terminal regions of the B-domain.

The amino acid numbering in the above construct is reproduced from the

factor VIII.

In order to index this example, the factor VIII:SQ amino acid

The amino acid numbering in the above construct is reproduced from the

fig, description in the specification. Note that Arg 740 is Arg 742

in the indexed sequence, etc. Also, Aen 745 (N-terminal link overlap)

in the map.

The amino acid numbering in the above construct is reproduced from the

factor VIII.

(Updated on 09-JAN-2003 to add missing OS field.)

(Updated on 25-MAR-2003 to correct PI field.)

Sequence 1440 AA.

Query Match 99.6%; Score 7659.5; DB 12; Length 1440;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1434; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

1 ATRTYVGLAVSLSDWNGSGLGELPVDARFPFPPKSPFFNPSVYKTLFVETVHLEN 60

2 ATRTYVGLAVSLSDWNGSGLGELPVDARFPFPPKSPFFNPSVYKTLFVETVHLEN 61

61 INKPPPPMGLGPTQAEYTVTVTLTKWASHPVSLHANGVYKASGSAVDYDQSQ 120

62 INKPPPPMGLGPTQAEYTVTVTLTKWASHPVSLHANGVYKASGSAVDYDQSQ 121

121 EKEDDVKVPGSHYTHYQVLENGPNSDPLCTVYLSHLDVLDNSGLCALICE 180

Db 122 RKEDKDVFGSGHTYVQVLKNGPMASDPCLCTLYSLVSLVDVLVGLNGLIGALLVCKR 181
 Qy 181 KESLAKECTQTLAKFTLITLAVFADSGLSMSBESFQSLKQGRDASLNAWPPQHTVNTYR 240
 Db 182 EESLAKENCTLAKFTLITLAVFADSGLSMSBESFQSLKQGRDASLNAWPPQHTVNTYR 241
 Qy 241 S-LPGLGCHESYVWNTIGWGTTPFVMSIFLEHGTFLVFRHQAASLIEPTFLPAQL 299
 Db 242 SLPGLGCHESYVWNTIGWGTTPFVMSIFLEHGTFLVFRHQAASLIEPTFLPAQL 301
 Qy 300 LMDQGLFLCHESQHDHCEAVYVYVQPEPLQMMNRESEYDDDLTDSBMDVR 359
 Db 302 LMDQGLFLCHESQHDHCEAVYVYVQPEPLQMMNRESEYDDDLTDSBMDVR 361
 Qy 360 FDDNSPSFQISVACKDCTVWYIAAEEDMDYAFVLPADRSYKSLANQPRI 419
 Db 362 FDDNSPSFQISVACKDCTVWYIAAEEDMDYAFVLPADRSYKSLANQPRI 421
 Qy 420 GRKYKVRPMAYDEFTFKREALQHESGILGPLAGVSGDGLAIPKQASRPNTPHG 479
 Db 422 GRKYKVRPMAYDEFTFKREALQHESGILGPLAGVSGDGLAIPKQASRPNTPHG 481
 Qy 480 ITDVPYERELKAGVGLHDPPLLAGHPKPKVTVVTDKPTSDPRCLTRYTSFVFN 539
 Db 482 ITDVPYERELKAGVGLHDPPLLAGHPKPKVTVVTDKPTSDPRCLTRYTSFVFN 541
 Qy 540 BRDASGLGLCYKESVDQGRQNSDKNSDKNSDKNSDKNSDKNSDKNSDKNSDKNSDK 599
 Db 542 BRDASGLGLCYKESVDQGRQNSDKNSDKNSDKNSDKNSDKNSDKNSDKNSDKNSDK 601
 Qy 600 GVLQDEDFPQSNMSTNGSVYFDSIQSLVCEHVAVTYLTLAGQDPLSVFSCYTK 659
 Db 602 GVLQDEDFPQSNMSTNGSVYFDSIQSLVCEHVAVTYLTLAGQDPLSVFSCYTK 661
 Qy 660 HKWYEDDTLTFPSGQTVFNKNGVGLNGLGNSDFNKNKGLVKKVSDQDQV 719
 Db 662 HKWYEDDTLTFPSGQTVFNKNGVGLNGLGNSDFNKNKGLVKKVSDQDQV 721
 Qy 720 FDSVDEYENLVGNNALEDRSPQNPVVRKHKQETVNTVNTVNTVNTVNTVNTVNTV 779
 Db 722 FDSVDEYENLVGNNALEDRSPQNPVVRKHKQETVNTVNTVNTVNTVNTVNTVNTV 781
 Qy 780 KEDPDTVDENKNSFQKTKRPTVAVELMDYKNSHSHVLNNAQSSQSPQFKV 839
 Db 782 KEDPDTVDENKNSFQKTKRPTVAVELMDYKNSHSHVLNNAQSSQSPQFKV 841
 Qy 840 VQRTFDGFTPLVRLNEHLGLGYTHAEVEDVNTFQKASRPTFSYSLISIE 899
 Db 842 VQRTFDGFTPLVRLNEHLGLGYTHAEVEDVNTFQKASRPTFSYSLISIE 901
 Qy 900 EDQRQAFPRNFVKNETTKYFKVQHPATDQEDFCANAFSDVLEKVSGLIG 959
 Db 902 EDQRQAFPRNFVKNETTKYFKVQHPATDQEDFCANAFSDVLEKVSGLIG 961
 Qy 960 PLVCHTINLPAHQGTQVDFLFTPTDETQTFNEMENKCAPCNQMDPTFK 1019
 Db 962 PLVCHTINLPAHQGTQVDFLFTPTDETQTFNEMENKCAPCNQMDPTFK 1021
 Qy 1020 ENVRFAHLMGTLEGLVAAQDQRWTLNLSGNSNHNHSFHSQVTVKKEEYEM 1079
 Db 1022 ENVRFAHLMGTLEGLVAAQDQRWTLNLSGNSNHNHSFHSQVTVKKEEYEM 1081
 Qy 1080 ALNVYFGVFTVPELFSKAGVWVCLCLAGELHAGSLTFLVYSNKKQTPQASGHR 1139
 Db 1082 ALNVYFGVFTVPELFSKAGVWVCLCLAGELHAGSLTFLVYSNKKQTPQASGHR 1141
 Qy 1140 DQITASGQGTQAPKLARHYSQSNMSTKPSFHWLVDLAPMIRKIQTKGQKQF 1199
 Db 1142 DQITASGQGTQAPKLARHYSQSNMSTKPSFHWLVDLAPMIRKIQTKGQKQF 1201
 Qy 1200 SSVLYTSQFTINSLDKQKWQTVKNSGTQVLAVFGVDSGCKNFTNPPI 1259
 Db 1202 SSVLYTSQFTINSLDKQKWQTVKNSGTQVLAVFGVDSGCKNFTNPPI 1261

Qy 1260 THYSIRSTRMLWELMGCDLNSCMFLGMSKISDAQITAGSIFTNMPTMSKSLHLQ 1319
 Db 1262 THYSIRSTRMLWELMGCDLNSCMFLGMSKISDAQITAGSIFTNMPTMSKSLHLQ 1321
 Qy 1320 GRSNAPVPPVAPKPEHLQVDFCKTKVTVGTVTQVTSLLLSHYVSEFLISSQDGHQTL 1379
 Db 1322 GRSNAPVPPVAPKPEHLQVDFCKTKVTVGTVTQVTSLLLSHYVSEFLISSQDGHQTL 1381
 Qy 1380 FPNQKQKVPQGNQDSFPPVNSLDPPLATVRLRHQSVHQTALRMBVLCGCAQDLY 1438
 Db 1382 FPNQKQKVPQGNQDSFPPVNSLDPPLATVRLRHQSVHQTALRMBVLCGCAQDLY 1440
 RESULT 8
 AAEL0832
 AC AAE10832 standard; Protein; 1459 AA.
 AC AAE10832;
 DT 18-DEC-2001 (first entry)
 DD Human factor VIII muten encoded by vector pUC8-2hyg-s.
 DD Human; hamostatic; coagulant; blood clotting factor; factor VIII;
 DD factor IX; therapy; haemophilia A; mutant; muten.
 DD Homo sapiens.
 DD Synthetic.
 XX Key
 XX Location/Qualifiers
 XX 19
 XX /label= Signal_peptide
 XX Protein
 XX 20..1459
 XX /product= "Human mature factor VIII muten"
 XX Region
 XX /label= Linker-peptide
 XX MO300170968-A2.
 XX 27-SEP-2001.
 XX
 XX 21-MAR-2001, 2001WO-RP03220.
 XX 22-MAR-2000, 2000EP-0106225.
 XX 08-MAY-2000, 2000US-203249F.
 XX (OCTA-) OCTAGENE GMBH.
 XX Hauser C, Hoerster A, Schroeder C, Lehnauer M;
 WPI; 2001-SS0175/66.
 N-PSDB; AAD48176.
 Recombinantly producing human blood coagulation factors VIII and IX for
 use in treating hemophilia -
 Claim 17; Pages 86-90; 10pp; English.
 The present invention relates to an improved method for the production
 of recombinant human blood clotting factors, especially factors VIII and
 IX, and to a method for the production of a recombinant human blood
 clotting factor, especially factor VIII, which is capable of expressing viral
 activators and carrying a genetic marker. The factor VIII muten or a gene
 encoding the blood coagulation factor. The factor VII muten or a gene
 transfer vector is used in the preparation of agents for treating or
 hemophilia, especially haemophilia A. The present sequence is
 identical to the sequence of the factor VIII muten. The factor VIII
 contains a silent mutation, resulting in a factor VII muten having
 the substitution of the B-domain of wild-type factor VIII by the
 linker peptide.
 Sequence 1459 AA;

Query Match		99.26;	Score 7626;	DB 22;	Length 1459;	
Similarity		99.26;	Score 7626;	DB 22;	Length 1459;	
Matches 1431;		Conservative	0;	Mismatches	3;	Indels 10;
						Gaps 2;
Qy	1	ATRYVLGAVELWQKQSDGELSDVADPRVPRVPSFPRVYVYKTLPISTWILFN	60			
Db	20	ATRYVLGAVELWQKQSDGELSDVADPRVPRVPSFPRVYVYKTLPISTWILFN	79			
Qy	61	TAKPRFHWGLLOFTQAEVDTVDITLQWASHPMSLAVGVSTWSEGAUDDQSQ	120			
Db	80	TAKPRFHWGLLOFTQAEVDTVDITLQWASHPMSLAVGVSTWSEGAUDDQSQ	139			
Qy	121	REKEDVYVPGGSSHTVQVLKENGPMASDPICLTISTSLGVTVQDLNGLGICALVCR	180			
Db	140	REKEDVYVPGGSSHTVQVLKENGPMASDPICLTISTSLGVTVQDLNGLGICALVCR	199			
Qy	181	BESLAKEKTLQALFVLAVFDGKSSMSSTNSLQDRDAASRAAMPPTVNYVVR	240			
Db	200	BESLAKEKTLQALFVLAVFDGKSSMSSTNSLQDRDAASRAAMPPTVNYVVR	259			
Qy	241	SILGLICRHSYVMVYVGMGTPTVPMISPLAGHTVTVNHRQASLEISPTTFATQLL	300			
Db	260	SILGLICRHSYVMVYVGMGTPTVPMISPLAGHTVTVNHRQASLEISPTTFATQLL	319			
Qy	301	MDGLGFLFCHISSEHQDQMAVYKVDCEPQLRMGNREADYDDUTDSBMDVRF	360			
Db	320	MDGLGFLFCHISSEHQDQMAVYKVDCEPQLRMGNREADYDDUTDSBMDVRF	379			
Qy	361	DDNDSFPTQIRESVAKKPTVHYTAAEEDMDYAPLAPDQDSYKSYQVLNPGQIG	420			
Db	380	DDNDSFPTQIRESVAKKPTVHYTAAEEDMDYAPLAPDQDSYKSYQVLNPGQIG	439			
Qy	421	KRYKVRPMYVDTETFKREAIOHESGLGPLYGVGGDTLALIFNQASRPYNIYHRI	480			
Db	440	KRYKVRPMYVDTETFKREAIOHESGLGPLYGVGGDTLALIFNQASRPYNIYHRI	499			
Qy	481	TDVRPLVSRRLPGVYHKLQFPLPGEIFKMYTVTVGGDTSPDLCTRYYSYVME	540			
Db	500	TDVRPLVSRRLPGVYHKLQFPLPGEIFKMYTVTVGGDTSPDLCTRYYSYVME	559			
Qy	541	RLASGLTGLPCLLYKSVSDRGVGTMDKRNVLVPSFDRNSKYLTVNIQRLNAG	600			
Db	560	RLASGLTGLPCLLYKSVSDRGVGTMDKRNVLVPSFDRNSKYLTVNIQRLNAG	619			
Qy	601	VLQGLGFPQALCYKSVSDRGVGTMDKRNVLVPSFDRNSKYLTVNIQRLNAG	660			
Db	620	VLQGLGFPQALCYKSVSDRGVGTMDKRNVLVPSFDRNSKYLTVNIQRLNAG	679			
Qy	661	KNVYEDTLTPFGSCVPSVNGVPMGLCGNSDNGVNIQALKSVQGVQDQVYH	720			
Db	680	KNVYEDTLTPFGSCVPSVNGVPMGLCGNSDNGVNIQALKSVQGVQDQVYH	739			
Qy	721	KNVYEDTLTPFGSCVPSVNGVPMGLCGNSDNGVNIQALKSVQGVQDQVYH	774			
Db	740	DSYEDLSAVLSNNALPERSFSON-----SRHQVYFZEGEITRTLTQSDESIDDDH	795			
Qy	775	SVWAKNDYTVYKRSQKSPQKQVTHVAVTEHVMQNSSPHVLNRAQSGSP	834			
Db	796	SVWAKNDYTVYKRSQKSPQKQVTHVAVTEHVMQNSSPHVLNRAQSGSP	855			
Qy	835	OPKVVQFQETDSTQPTLREGLNEHGLGSPITRAVTEHVMQNSSPHVLNRAQSGSP	894			
Db	856	OPKVVQFQETDSTQPTLREGLNEHGLGSPITRAVTEHVMQNSSPHVLNRAQSGSP	915			
Qy	895	LSTYEDQCAEPKQVPMENKTYPMQVQHPMAFDQEPCKAMAFSDVDLQDVR	954			
Db	916	LSTYEDQCAEPKQVPMENKTYPMQVQHPMAFDQEPCKAMAFSDVDLQDVR	975			
Qy	955	SGLIGPLVCTVTLNPHRGQVTVQVPLFTFDSTSSYTFNNKRCRACQIOME	1014			
Db	976	SGLIGPLVCTVTLNPHRGQVTVQVPLFTFDSTSSYTFNNKRCRACQIOME	1035			
Qy	1015	DTDFKTEKREFAINTIDMTLFGLVQAQDRTWYLLSGVSSNENHSINFSGRVYVKK	1074			
Db	1036	DTDFKTEKREFAINTIDMTLFGLVQAQDRTWYLLSGVSSNENHSINFSGRVYVKK	1095			
Qy	1075	BEYKALNLYITPVFFTEALFSAGIWEVCIIGELHAGAGSTLFLVSNKCTPLGAA	1134			
Db	1096	BEYKALNLYITPVFFTEALFSAGIWEVCIIGELHAGAGSTLFLVSNKCTPLGAA	1155			
Qy	1135	SGHIDEDQITASSGVQQAAPLARLHYSINANSTKPFSSIKVOLLAPMIHGIKTG	1194			
Db	1156	SGHIDEDQITASSGVQQAAPLARLHYSINANSTKPFSSIKVOLLAPMIHGIKTG	1215			
Qy	1195	ARQKPSSELYISQFIINYSLQKQVQVGNSTQTLVTFGVNDSGGIKHMIENPIYAK	1254			
Db	1216	ARQKPSSELYISQFIINYSLQKQVQVGNSTQTLVTFGVNDSGGIKHMIENPIYAK	1275			
Qy	1255	RLHPTHTSIRSTLHEMLQCDLSCMPJLHMSKALSDQITASSYTFNNFATMSPSA	1314			
Db	1276	RLHPTHTSIRSTLHEMLQCDLSCMPJLHMSKALSDQITASSYTFNNFATMSPSA	1335			
Qy	1315	RLHPTHTSIRSTLHEMLQCDLSCMPJLHMSKALSDQITASSYTFNNFATMSPSA	1374			
Db	1336	RLHPTHTSIRSTLHEMLQCDLSCMPJLHMSKALSDQITASSYTFNNFATMSPSA	1395			
Qy	1375	QVHTLPPQKQVYVQCDQSTFPPVNSLDPLFLVLRHPSQVHQAALNEMVLCGA	1434			
Db	1396	QVHTLPPQKQVYVQCDQSTFPPVNSLDPLFLVLRHPSQVHQAALNEMVLCGA	1455			
Qy	1435	ODX 1438				
Db	1456	ODX 1459				
RESULT 9		AAEI0833 standard; Protein; 1459 AA.				
ID		AAEI0833				
AC		AA				
DC		AA				
DT		18-DEC-2001 (first entry)				
XX		Human factor VIII protein encoded by vector p078-3.				
XX		Human; haemostatic; coagulant; blood clotting factor; factor VIII;				
XX		factor IX; therapy; haemophilia A; mutant; mucin.				
XX		Homo sapiens.				
XX		Synthetic.				
XX		Key				
XX		Peptide				
XX		Location/Qualifiers				
XX		/label= signal_peptide				
XX		Protein				
XX		Product= "Human mature factor VIII mucin"				
XX		Misc-difference				
XX		/note= "Wild type Val substituted with Ala"				
XX		Region				
XX		/label= linker-peptide				
XX		Misc-difference				
XX		/note= "Wild type Val substituted with Glu"				
XX		ID				
XX		27-SEP-2001.				
XX		21-MAR-2001; 2001WO-BP03220.				
XX		22-MAR-2000; 2000EP-0196295.				
XX		08-MAY-2000; 2000US-203249F.				
XX		(OCTA-) OCTAGINE GMBH.				
XX		Hauser C, Hoerster A, Schroeder C, Lehnauer M;				
XX						

1322 NIENPIATRYAPPTVSTJAMELWCLNCSNPLQWESKASDNATASVPT 1381
 1304 INPNTNPFKARILHQSSNARVPOVNPFFLOVDQCTKMYGTQVTVQVSLSNVY 1363
 1382 INVAIIVKSSKARILHQSSNARVPOVNPFFLOVDQCTKMYGTQVTVQVSLSNVY 1441
 1364 KEPLISSQDCHQNTLFTQNGKTVVPOQSDSTFVNSLQPLATRYLRHPQSWHOI 1423
 1442 KEPLISSQDCHQNTLFTQNGKTVVPOQSDSTFVNSLQPLATRYLRHPQSWHOI 1501
 1424 ALRMVTEGCEQVLY 1438
 1502 AUGRGTGCEADQMT 1516
 1504
 1506
 1508
 1510
 1512
 1514
 1516
 1518
 1520
 1522
 1524
 1526
 1528
 1530
 1532
 1534
 1536
 1538
 1540
 1542
 1544
 1546
 1548
 1550
 1552
 1554
 1556
 1558
 1560
 1562
 1564
 1566
 1568
 1570
 1572
 1574
 1576
 1578
 1580
 1582
 1584
 1586
 1588
 1590
 1592
 1594
 1596
 1598
 1600
 1602
 1604
 1606
 1608
 1610
 1612
 1614
 1616
 1618
 1620
 1622
 1624
 1626
 1628
 1630
 1632
 1634
 1636
 1638
 1640
 1642
 1644
 1646
 1648
 1650
 1652
 1654
 1656
 1658
 1660
 1662
 1664
 1666
 1668
 1670
 1672
 1674
 1676
 1678
 1680
 1682
 1684
 1686
 1688
 1690
 1692
 1694
 1696
 1698
 1700
 1702
 1704
 1706
 1708
 1710
 1712
 1714
 1716
 1718
 1720
 1722
 1724
 1726
 1728
 1730
 1732
 1734
 1736
 1738
 1740
 1742
 1744
 1746
 1748
 1750
 1752
 1754
 1756
 1758
 1760
 1762
 1764
 1766
 1768
 1770
 1772
 1774
 1776
 1778
 1780
 1782
 1784
 1786
 1788
 1790
 1792
 1794
 1796
 1798
 1800
 1802
 1804
 1806
 1808
 1810
 1812
 1814
 1816
 1818
 1820
 1822
 1824
 1826
 1828
 1830
 1832
 1834
 1836
 1838
 1840
 1842
 1844
 1846
 1848
 1850
 1852
 1854
 1856
 1858
 1860
 1862
 1864
 1866
 1868
 1870
 1872
 1874
 1876
 1878
 1880
 1882
 1884
 1886
 1888
 1890
 1892
 1894
 1896
 1898
 1900
 1902
 1904
 1906
 1908
 1910
 1912
 1914
 1916
 1918
 1920
 1922
 1924
 1926
 1928
 1930
 1932
 1934
 1936
 1938
 1940
 1942
 1944
 1946
 1948
 1950
 1952
 1954
 1956
 1958
 1960
 1962
 1964
 1966
 1968
 1970
 1972
 1974
 1976
 1978
 1980
 1982
 1984
 1986
 1988
 1990
 1992
 1994
 1996
 1998
 2000
 2002
 2004
 2006
 2008
 2010
 2012
 2014
 2016
 2018
 2020
 2022
 2024
 2026
 2028
 2030
 2032
 2034
 2036
 2038
 2040
 2042
 2044
 2046
 2048
 2050
 2052
 2054
 2056
 2058
 2060
 2062
 2064
 2066
 2068
 2070
 2072
 2074
 2076
 2078
 2080
 2082
 2084
 2086
 2088
 2090
 2092
 2094
 2096
 2098
 2100
 2102
 2104
 2106
 2108
 2110
 2112
 2114
 2116
 2118
 2120
 2122
 2124
 2126
 2128
 2130
 2132
 2134
 2136
 2138
 2140
 2142
 2144
 2146
 2148
 2150
 2152
 2154
 2156
 2158
 2160
 2162
 2164
 2166
 2168
 2170
 2172
 2174
 2176
 2178
 2180
 2182
 2184
 2186
 2188
 2190
 2192
 2194
 2196
 2198
 2200
 2202
 2204
 2206
 2208
 2210
 2212
 2214
 2216
 2218
 2220
 2222
 2224
 2226
 2228
 2230
 2232
 2234
 2236
 2238
 2240
 2242
 2244
 2246
 2248
 2250
 2252
 2254
 2256
 2258
 2260
 2262
 2264
 2266
 2268
 2270
 2272
 2274
 2276
 2278
 2280
 2282
 2284
 2286
 2288
 2290
 2292
 2294
 2296
 2298
 2300
 2302
 2304
 2306
 2308
 2310
 2312
 2314
 2316
 2318
 2320
 2322
 2324
 2326
 2328
 2330
 2332
 2334
 2336
 2338
 2340
 2342
 2344
 2346
 2348
 2350
 2352
 2354
 2356
 2358
 2360
 2362
 2364
 2366
 2368
 2370
 2372
 2374
 2376
 2378
 2380
 2382
 2384
 2386
 2388
 2390
 2392
 2394
 2396
 2398
 2400
 2402
 2404
 2406
 2408
 2410
 2412
 2414
 2416
 2418
 2420
 2422
 2424
 2426
 2428
 2430
 2432
 2434
 2436
 2438
 2440
 2442
 2444
 2446
 2448
 2450
 2452
 2454
 2456
 2458
 2460
 2462
 2464
 2466
 2468
 2470
 2472
 2474
 2476
 2478
 2480
 2482
 2484
 2486
 2488
 2490
 2492
 2494
 2496
 2498
 2500
 2502
 2504
 2506
 2508
 2510
 2512
 2514
 2516
 2518
 2520
 2522
 2524
 2526
 2528
 2530
 2532
 2534
 2536
 2538
 2540
 2542
 2544
 2546
 2548
 2550
 2552
 2554
 2556
 2558
 2560
 2562
 2564
 2566
 2568
 2570
 2572
 2574
 2576
 2578
 2580
 2582
 2584
 2586
 2588
 2590
 2592
 2594
 2596
 2598
 2600
 2602
 2604
 2606
 2608
 2610
 2612
 2614
 2616
 2618
 2620
 2622
 2624
 2626
 2628
 2630
 2632
 2634
 2636
 2638
 2640
 2642
 2644
 2646
 2648
 2650
 2652
 2654
 2656
 2658
 2660
 2662
 2664
 2666
 2668
 2670
 2672
 2674
 2676
 2678
 2680
 2682
 2684
 2686
 2688
 2690
 2692
 2694
 2696
 2698
 2700
 2702
 2704
 2706
 2708
 2710
 2712
 2714
 2716
 2718
 2720
 2722
 2724
 2726
 2728
 2730
 2732
 2734
 2736
 2738
 2740
 2742
 2744
 2746
 2748
 2750
 2752
 2754
 2756
 2758
 2760
 2762
 2764
 2766
 2768
 2770
 2772
 2774
 2776
 2778
 2780
 2782
 2784
 2786
 2788
 2790
 2792
 2794
 2796
 2798
 2800
 2802
 2804
 2806
 2808
 2810
 2812
 2814
 2816
 2818
 2820
 2822
 2824
 2826
 2828
 2830
 2832
 2834
 2836
 2838
 2840
 2842
 2844
 2846
 2848
 2850
 2852
 2854
 2856
 2858
 2860
 2862
 2864
 2866
 2868
 2870
 2872
 2874
 2876
 2878
 2880
 2882
 2884
 2886
 2888
 2890
 2892
 2894
 2896
 2898
 2900
 2902
 2904
 2906
 2908
 2910
 2912
 2914
 2916
 2918
 2920
 2922
 2924
 2926
 2928
 2930
 2932
 2934
 2936
 2938
 2940
 2942
 2944
 2946
 2948
 2950
 2952
 2954
 2956
 2958
 2960
 2962
 2964
 2966
 2968
 2970
 2972
 2974
 2976
 2978
 2980
 2982
 2984
 2986
 2988
 2990
 2992
 2994
 2996
 2998
 3000
 3002
 3004
 3006
 3008
 3010
 3012
 3014
 3016
 3018
 3020
 3022
 3024
 3026
 3028
 3030
 3032
 3034
 3036
 3038
 3040
 3042
 3044
 3046
 3048
 3050
 3052
 3054
 3056
 3058
 3060
 3062
 3064
 3066
 3068
 3070
 3072
 3074
 3076
 3078
 3080
 3082
 3084
 3086
 3088
 3090
 3092
 3094
 3096
 3098
 3100
 3102
 3104
 3106
 3108
 3110
 3112
 3114
 3116
 3118
 3120
 3122
 3124
 3126
 3128
 3130
 3132
 3134
 3136
 3138
 3140
 3142
 3144
 3146
 3148
 3150
 3152
 3154
 3156
 3158
 3160
 3162
 3164
 3166
 3168
 3170
 3172
 3174
 3176
 3178
 3180
 3182
 3184
 3186
 3188
 3190
 3192
 3194
 3196
 3198
 3200
 3202
 3204
 3206
 3208
 3210
 3212
 3214
 3216
 3218
 3220
 3222
 3224
 3226
 3228
 3230
 3232
 3234
 3236
 3238
 3240
 3242
 3244
 3246
 3248
 3250
 3252
 3254
 3256
 3258
 3260
 3262
 3264
 3266
 3268
 3270
 3272
 3274
 3276
 3278
 3280
 3282
 3284
 3286
 3288
 3290
 3292
 3294
 3296
 3298
 3300
 3302
 3304
 3306
 3308
 3310
 3312
 3314
 3316
 3318
 3320
 3322
 3324
 3326
 3328
 3330
 3332
 3334
 3336
 3338
 3340
 3342
 3344
 3346
 3348
 3350
 3352
 3354
 3356
 3358
 3360
 3362
 3364
 3366
 3368
 3370
 3372
 3374
 3376
 3378
 3380
 3382
 3384
 3386
 3388
 3390
 3392
 3394
 3396
 3398
 3400
 3402
 3404
 3406
 3408
 3410
 3412
 3414
 3416
 3418
 3420
 3422
 3424
 3426
 3428
 3430
 3432
 3434
 3436
 3438
 3440
 3442
 3444
 3446
 3448
 3450
 3452
 3454
 3456
 3458
 3460
 3462
 3464
 3466
 3468
 3470
 3472
 3474
 3476
 3478
 3480
 3482
 3484
 3486
 3488
 3490
 3492
 3494
 3496
 3498
 3500
 3502
 3504
 3506
 3508
 3510
 3512
 3514
 3516
 3518
 3520
 3522
 3524
 3526
 3528
 3530
 3532
 3534
 3536
 3538
 3540
 3542
 3544
 3546
 3548
 3550
 3552
 3554
 3556
 3558
 3560
 3562
 3564
 3566
 3568
 3570
 3572
 3574
 3576
 3578
 3580
 3582
 3584
 3586
 3588
 3590
 3592
 3594
 3596
 3598
 3600
 3602
 3604
 3606
 3608
 3610
 3612
 3614
 3616
 3618
 3620
 3622
 3624
 3626
 3628
 3630
 3632
 3634
 3636
 3638
 3640
 3642
 3644
 3646
 3648
 3650
 3652
 3654
 3656
 3658
 3660
 3662
 3664
 3666
 3668
 3670
 3672
 3674
 3676
 3678
 3680
 3682
 3684
 3686
 3688
 3690
 3692
 3694
 3696
 3698
 3700
 3702
 3704
 3706
 3708
 3710
 3712
 3714
 3716
 3718
 3720
 3722
 3724
 3726
 3728
 3730
 3732
 3734
 3736
 3738
 3740
 3742
 3744
 3746
 3748
 3750
 3752
 3754
 3756
 3758
 3760
 3762
 3764
 3766
 3768
 3770
 3772
 3774
 3776
 3778
 3780
 3782
 3784
 3786
 3788
 3790
 3792
 3794
 3796
 3798
 3800
 3802
 3804
 3806
 3808
 3810
 3812
 3814
 3816
 3818
 3820
 3822
 3824
 3826
 3828
 3830
 3832
 3834
 3836
 3838
 3840
 3842
 3844
 3846
 3848
 3850
 3852
 3854
 3856
 3858
 3860
 3862
 3864
 3866
 3868
 3870
 3872
 3874
 3876
 3878
 3880
 3882
 3884
 3886
 3888
 3890
 3892
 3894
 3896
 3898
 3900
 3902
 3904
 3906
 3908
 3910
 3912
 3914
 3916
 3918
 3920
 3922
 3924
 3926
 3928
 3930
 3932
 3934
 3936
 3938
 3940
 3942
 3944
 3946
 3948
 3950
 3952
 3954
 3956
 3958
 3960
 3962
 3964
 3966
 3968
 3970
 3972
 3974
 3976
 3978
 3980
 3982
 3984
 3986
 3988
 3990
 3992
 3994
 3996
 3998
 4000
 4002
 4004
 4006
 4008
 4010
 4012
 4014
 4016
 4018
 4020
 4022
 4024
 4026
 4028
 4030
 4032
 4034
 4036
 4038
 4040
 4042
 4044
 4046
 4048
 4050
 4052
 4054
 4056
 4058
 4060
 4062
 4064
 4066
 4068
 4070
 4072
 4074
 4076
 4078
 4080
 4082
 4084
 4086
 4088
 4090
 4092
 4094
 4096
 4098
 4100
 4102
 4104
 4106
 4108
 4110
 4112
 4114
 4116
 4118
 4120
 4122
 4124
 4126
 4128
 4130
 4132
 4134
 4136
 4138
 4140
 4142
 4144
 4146
 4148
 4150
 4152
 4154
 4156
 4158

QY 1021 NYRPHAINGTIMTLPGLMAQDQRIMYLKMGKSNENIHFSHQVTVRRKBYDQA 1080
 Db 1007 NYRPHAINGTIMTLPGLMAQDQRIMYLKMGKSNENIHFSHQVTVRRKBYDQA 1066
 QY 1081 LNYLVEGVETVSMRPAKGIWVRVCLGHEHJHAGNGLFLVYSKCPQLQWAGSHFED 1140
 Db 1067 LNYLVEGVETVSMRPAKGIWVRVCLGHEHJHAGNGLFLVYSKCPQLQWAGSHFED 1126
 QY 1141 FQTTASGQGVNAPKALAHYSGSINAMSTPEPSWIKVLDLAPMLHIGLTKCARQKES 1200
 Db 1127 FQTTASGQGVNAPKALAHYSGSINAMSTPEPSWIKVLDLAPMLHIGLTKCARQKES 1186
 QY 1201 SIXTISQRTIMYSLQCKPQVYGNSTGTLAVFQNSGSSLENNENPRLIARVILHPT 1260
 Db 1187 SIXTISQRTIMYSLQCKPQVYGNSTGTLAVFQNSGSSLENNENPRLIARVILHPT 1246
 QY 1261 HYSITRSTLWELWGCGLNCSNPLCHESKLSIDQAIQASSTVFWFWRQSPKARHJGQ 1320
 Db 1247 HYSITRSTLWELWGCGLNCSNPLCHESKLSIDQAIQASSTVFWFWRQSPKARHJGQ 1306
 QY 1321 RSNARNPQVNNKQKQWQKQWQKQWQKQWQKQWQKQWQKQWQKQWQKQWQKQW 1380
 Db 1307 RSNARNPQVNNKQKQWQKQWQKQWQKQWQKQWQKQWQKQWQKQWQKQWQKQW 1366
 QY 1361 FQNGKVKVFGNQDQSFPPVAVSLQPLPLTYLIRHQSGVQVQLNLRVLEAQQYQL 1424
 Db 1367 FQNGKVKVFGNQDQSFPPVAVSLQPLPLTYLIRHQSGVQVQLNLRVLEAQQYQL 1424

RESULT 13

AA018622
 ID AA018622 standard; Protein; 1424 AA.

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

XX XX
 XX XX
 XX XX

CC reduced sulfate proteoglycan (HSPG)-dependent or receptor-independent
 CC clearance and procollagen activity, which has a nonconservative amino
 CC acid substitution at one or more positions in the A2 domain consisting of
 CC 491 (490/491) or Arg(490/491) or in the C2
 CC domain relative to the wild-type VII. The VII
 CC polypeptide encoding it and a receptor-associated protein (alpha2
 CC macroglobulin receptor-associated protein or RAP) are useful for treating
 CC macrophilia. The mutated protein has a longer half-life. The present
 CC sequence is the human mature B-domainless factor VIII protein.
 XX Sequence 1424 AA;

Query Match 98.7%; Score 7592; DB 23; Length 1424;

Best Local Similarity 99.0%; Pos. No. 0;

Matches 1423; Conservative 1; Mismatches 1; Indels 14; Gaps 1;

QY 1 ATNRYVGLVSLSDMSLDELVDLQFPPVPSFSPFSTVYKTLVPTVHLFNF 60
 Db 1 ATNRYVGLVSLSDMSLDELVDLQFPPVPSFSPFSTVYKTLVPTVHLFNF 60
 QY 61 TADSPFPPHGLGPTTQATVTVTVTLQWASHPVLAQVGVSTWASAGAEVDDQSQ 120
 Db 61 TADSPFPPHGLGPTTQATVTVTVTLQWASHPVLAQVGVSTWASAGAEVDDQSQ 120
 QY 121 REKEDDVFVPGSHRTVWOLKNGPMASDDCLCTSTSLHVDLVKLNGLICALLVCR 180
 Db 121 REKEDDVFVPGSHRTVWOLKNGPMASDDCLCTSTSLHVDLVKLNGLICALLVCR 180
 QY 181 EQLAKETQTLKFLTFVAFDQSGHSHSTENLAQEDASANRPMPTVNGTVNR 240
 Db 181 EQLAKETQTLKFLTFVAFDQSGHSHSTENLAQEDASANRPMPTVNGTVNR 240
 QY 241 SLPLTGCHRSYTHVVGTTPEVHSIFLEGHTELVNHRQASLEISPTITLQTL 300
 Db 241 SLPLTGCHRSYTHVVGTTPEVHSIFLEGHTELVNHRQASLEISPTITLQTL 300
 QY 301 MDLGGPFLFCHSSHQDGEAVYKVSQPEPQLMKHNEAEVDYDDLTDSMDVVRP 360
 Db 301 MDLGGPFLFCHSSHQDGEAVYKVSQPEPQLMKHNEAEVDYDDLTDSMDVVRP 360
 QY 361 DDNSPSFTQHSVAKSHKPTWHVYLAEEEDNDVAVLAPDRYSQYLANGPQIG 420
 Db 361 DDNSPSFTQHSVAKSHKPTWHVYLAEEEDNDVAVLAPDRYSQYLANGPQIG 420
 QY 421 KYKVEVPMATDPTETRAALQHSGLGQKVEQVGLTLFKNQASRHYVPHGI 480
 Db 421 KYKVEVPMATDPTETRAALQHSGLGQKVEQVGLTLFKNQASRHYVPHGI 480
 QY 481 TDVRYSLRPLKGVOLQKDFPLPQRTFKYKVTYEDQPKDQCLVYSPGRN 540
 Db 481 TDVRYSLRPLKGVOLQKDFPLPQRTFKYKVTYEDQPKDQCLVYSPGRN 540
 QY 541 DLALAGLGLPLCTHESVQKQNLNSYVLPGRNRYHTFEMTQFLDNPAG 600
 Db 541 DLALAGLGLPLCTHESVQKQNLNSYVLPGRNRYHTFEMTQFLDNPAG 600
 QY 601 VQLEPFRQASVKNRYNTQVYSQSLVCLRPVAVYLLTGAQDTLSVPSGTFPH 660
 Db 601 VQLEPFRQASVKNRYNTQVYSQSLVCLRPVAVYLLTGAQDTLSVPSGTFPH 660
 QY 661 KVVVDVTLTFPFFSTFVSNFQMLTLCQNSDPFNRGATALLKYSQCDQKDYTE 720
 Db 661 KVVVDVTLTFPFFSTFVSNFQMLTLCQNSDPFNRGATALLKYSQCDQKDYTE 720
 QY 721 DSYDTSATLSEKRNALPSPRSQPPVAKSHQRIETATTLQSQEELIDDTLVNKK 780
 Db 721 DSYDTSATLSEKRNALPSPRSQPPVAKSHQRIETATTLQSQEELIDDTLVNKK 780
 QY 781 KDDTVDDENGSPFQKTRHTFYATVELMDYQSSSFVLRNQAQSGVDFKCV 840
 Db 781 KDDTVDDENGSPFQKTRHTFYATVELMDYQSSSFVLRNQAQSGVDFKCV 840
 QY 841 FQETDGSFTQPLKREINHEHLLGPIYRAEVENIMVTFNQASRPYSFVSLSTE 900
 Db 841 FQETDGSFTQPLKREINHEHLLGPIYRAEVENIMVTFNQASRPYSFVSLSTE 900

XX Human mature B-domainless factor VIII.
 XX Human; factor VIII; fVIII; half-life; mutant; haemophilia;
 XX heparan sulfate proteoglycan-mediated clearance; RAP;
 XX receptor-associated protein; haemoretic; gene therapy;
 XX alpha2 macroglobulin receptor-associated protein.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 XX Domain 373..740
 XX /label= A2_domain
 XX Region 494..509
 XX /label= LRP_binding_region
 XX W0200260951-A2.
 XX 08-AUG-2002.
 XX 11-JAN-2002; 2002WO-US00583.
 XX 12-JAN-2001; 2001US-260904P.
 XX (JANNA-) AMERICAN NAT RED CROSS.
 XX Saenko EL, Sarafanov AG;
 XX WPI; 2002-60850/65.
 XX New mutant factor VIII with reduced sulfate proteoglycan
 XX (HSPG)-dependent or receptor-independent clearance and procollagen
 XX activity for treating hemophilia -
 XX Claim 9; Fig 12; 161pp; English.
 XX The present invention relates to a mutant factor VIII protein with

Db 827 FORTTASFTQPLVGEHLLGGLGPTVRAENVNWTTPQASPPVSYSLSTYS 886
 Qy 901 DORGGPRAPFTVNTNTYTKVQHHMAFKDQPCDANAFSDVLEKDVHSLGIP 960
 Db 887 DQGGKPRKFNFTVNTNTYTKVQHHMAFKDQPCDANAFSDVLEKDVHSLGIP 946
 Qy 961 LLVCHNTNLPAGGQVTFQEPALFTPTDQKSNFTFENMRNCRAPQDQEDPTKS 1020
 Db 947 LQVCHNTNLPAGGQVTFQEPALFTPTDQKSNFTFENMRNCRAPQDQEDPTKS 1006
 Qy 1021 NTRPHALNGTMTPLGMAQDQRIWILLKSNENHSHFSGHVTYKKEBYTQA 1080
 Db 1007 NTRPHALNGTMTPLGMAQDQRIWILLKSNENHSHFSGHVTYKKEBYTQA 1066
 Qy 1081 KYRFLYFGVETVELMPSKAGVRYECLIGHHAGNSHTLFVYNSKQPLGMAQSHIRO 1140
 Db 1067 KYRFLYFGVETVELMPSKAGVRYECLIGHHAGNSHTLFVYNSKQPLGMAQSHIRO 1126
 Qy 1141 FOITASQGVQWAPKLARHVSIGSNAMSTKPSFKVDGLAPMLHIGIKTGAKQKFS 1200
 Db 1127 FOITASQGVQWAPKLARHVSIGSNAMSTKPSFKVDGLAPMLHIGIKTGAKQKFS 1186
 Qy 1201 SIXYISQELIMSLDQKQVTVGNSGTGLAMVFGNVDSQTKNTNPPPIIARYIRHPT 1260
 Db 1187 SIXYISQELIMSLDQKQVTVGNSGTGLAMVFGNVDSQTKNTNPPPIIARYIRHPT 1246
 Qy 1261 HYSRSLRLEWELMGCLNCSPLMGESKAIIDAQVTAASFYFNPAFWPSKARHLQ 1320
 Db 1247 HYSRSLRLEWELMGCLNCSPLMGESKAIIDAQVTAASFYFNPAFWPSKARHLQ 1306
 Qy 1321 RSNMRPQNNPKZWLOVDFQTKVQVTVGVNSLLTSMTYKEFLSSQDGHQMTLF 1366
 Db 1307 RSNMRPQNNPKZWLOVDFQTKVQVTVGVNSLLTSMTYKEFLSSQDGHQMTLF 1366
 Qy 1361 FONGKVKYFGQNDQSFPPVNSLDPPLLTYLRHQSHVQVHQLRWLVGLCEADY 1438
 Db 1367 FONGKVKYFGQNDQSFPPVNSLDPPLLTYLRHQSHVQVHQLRWLVGLCEADY 1424

RESULT 14

ABG92541 standard; Protein; 1447 AA.
 ID ABG92541
 AC ABG92541

XX 19-NOV-2002 (first entry)
 XX XX
 XX XX

XX SAg3 B-domain-deleted-Factor VIII (FVIII).
 XX XX

XX Human, alpha-galactosidase; Factor VIII; Factor IX;
 XX XX

XX Homo sapiens.
 XX XX

XX Synthetic.
 XX XX

XX W0200241799-A2.
 XX XX

XX 22-AUG-2002.
 XX XX

XX 11-OCT-2001; 2001NC-0542655.
 XX XX

XX 11-OCT-2000; 2000US-0686497.
 XX XX

XX (TRAN-) TRANSKARYOTIC THERAPIES INC.
 XX XX

XX Seldon RF, Miller AM, Treco DS;
 XX XX

XX MP2; 2002-627600/67.
 XX XX

XX A-75007; A5864006.
 XX XX

XX New synthetic nucleic acid sequence involving a continuous stretch of
 XX XX

XX At least 150 common codons, useful for expressing mammalian, preferable
 XX XX

human proteins e.g. alpha-galactosidase or Factor VIII or IX or for
 gene therapy -

Example 1; Figure 9; 115pp; English.

The invention relates to a synthetic nucleic acid sequence, where at
 least one non-common or less-common codon is replaced with a common
 codon. The synthetic nucleic acid is useful for expressing mammalian,
 preferably human proteins e.g. alpha-galactosidase or Factor VIII or IX
 for gene therapy. The synthetic nucleic acid allows precise dosing and
 reduces treatment cost. The invention is useful for gene therapy and
 is curative (one gene therapy treatment has the potential to last a
 patient's lifetime. ABG92540-ABG92541 represent Factor VIII and Factor
 IX synthetic amino acid sequences of the invention.

Sequence 1447 AA;

Query Match 98.7%; Score 7588; DB 23; Length 1447;

Similarity 98.7%; Score 7588; DB 23; Length 1447;

Matches 1423; Conservative 1; Mismatches 4; Indels 10; Gaps 1;

1 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

2 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

3 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

4 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

5 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

6 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

7 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

8 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

9 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

10 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

11 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

12 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

13 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

14 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

15 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

16 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

17 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

18 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

19 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

20 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

21 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

22 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

23 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

24 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

25 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

26 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

27 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

28 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

29 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

30 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

31 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

32 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

33 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

34 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

35 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

36 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

37 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

38 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

39 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

40 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

41 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

42 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

43 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

44 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

45 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

46 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

47 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

48 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

49 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

50 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

51 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

52 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

53 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

54 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

55 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

56 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

57 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

58 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

59 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

60 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

61 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

62 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

63 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

64 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

65 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

66 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

67 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

68 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

69 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

70 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

71 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

72 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

73 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

74 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

75 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

76 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

77 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

78 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

79 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

80 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

81 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

82 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

83 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

84 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

85 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

86 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

87 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

88 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

89 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

90 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

91 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

92 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

93 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

94 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

95 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

96 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

97 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

98 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

99 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

100 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

101 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

102 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

103 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

104 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

105 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

106 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

107 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

108 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

109 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

110 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

111 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

112 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

113 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

114 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

115 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

116 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

117 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

118 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

119 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

120 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

121 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

122 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

123 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

124 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

125 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

126 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

127 ATRATYGLAVSLMDMDSGLGSELPVDAVPPVPPKPPFPSTVYKTKLTFVETVHLRN 60

128 ATRATYGLAVSLMDMDSGLG

QY 781 EDPFDYEDENQEPSPQKTRHYFAAVERLMDYMGSSPHWLRNRAQSGVQPKKV 840
 Db 790 EDPFDYEDENQEPSPQKTRHYFAAVERLMDYMGSSPHWLRNRAQSGVQPKKV 849
 QY 841 FOEPTDQSGTFOPIRGELNHELGLLOPIRAEVDENIHWTFEKNQSRPVSYSLLSYEE 900
 Db 850 FOEPTDQSGTFOPIRGELNHELGLLOPIRAEVDENIHWTFEKNQSRPVSYSLLSYEE 909
 QY 901 QOEGAGREPRFQVNETKYFWKVOHPAPWADKDFEPCAKWATSDVLEKQVHSLGP 960
 Db 910 QOEGAGREPRFQVNETKYFWKVOHPAPWADKDFEPCAKWATSDVLEKQVHSLGP 969
 QY 961 LAMCHWTINLWENSGOVTVQEPALFTTIDETKSATVTFENMEKACQNIQWDEPFKE 1020
 Db 970 LAMCHWTINLWENSGOVTVQEPALFTTIDETKSATVTFENMEKACQNIQWDEPFKE 1029
 QY 1021 MYRPHANQYVWTLQGLWADQDQRTWLLYSGNSNTHTSHGVTFFVAKSEYKKA 1080
 Db 1030 MYRPHANQYVWTLQGLWADQDQRTWLLYSGNSNTHTSHGVTFFVAKSEYKKA 1089
 QY 1081 LYNLWVQVETFWMLPSKAGTIVUQGLKQHLHAGKQETLFTYKNGCOTPYQMAQSHRD 1140
 Db 1090 LYNLWVQVETFWMLPSKAGTIVUQGLKQHLHAGKQETLFTYKNGCOTPYQMAQSHRD 1149
 QY 1141 PQTASQGVQWAKARLWQSGIYNWENRPSHWKTLAAHMHGHTIQWQAKQFS 1200
 Db 1150 PQTASQGVQWAKARLWQSGIYNWENRPSHWKTLAAHMHGHTIQWQAKQFS 1209
 QY 1201 GUYTGQITWMLGQWQVTKRNGSTQWVYFQWDSQSTKQNTFNPPIIARTILHPT 1260
 Db 1210 GUYTGQITWMLGQWQVTKRNGSTQWVYFQWDSQSTKQNTFNPPIIARTILHPT 1269
 QY 1261 HYSIRLSTLWELADJLWCSWPLQWQSHSAISDAQITASSTFTFATWSKARLHJQ 1320
 Db 1270 HYSIRLSTLWELADJLWCSWPLQWQSHSAISDAQITASSTFTFATWSKARLHJQ 1329
 QY 1321 SSNAWEPQVWENLQVDFQWTKWQVQVQVQKSLTATWVWKEPLTSSQDQWQTLF 1380
 Db 1330 SSNAWEPQVWENLQVDFQWTKWQVQVQKSLTATWVWKEPLTSSQDQWQTLF 1389
 QY 1381 FQNGKVVQFQNSQSTFWVNSLQPLRYLRLZHPQSHVQWQALQWVGEQWQLY 1438
 Db 1390 FQNGKVVQFQNSQSTFWVNSLQPLRYLRLZHPQSHVQWQALQWVGEQWQLY 1447

RESULT 15

1D AAR60267 standard; protein; 1425 AA.

AC AA AAR60267;

XX 25-MAR-2003 (updated)

DT 10-OCT-1990 (first entry)

XX Modified factor VIII:C sequence with the R740-E1649 deletion.

XX Modified factor VIII:C; maturation polypeptide; haemophilias;
 KW blood coagulation; RE deletion.

XX Homo sapiens.

XX W08080831-A.

XX 11-FEB-1988.

XX 31-JUL-1987; 87WO-US01814.

XX 01-AUG-1986; 86US-0899375.

XX (BIQ) BIOGEN NV.

PA (PASE) PASEK M P.

XX Pasek MF;

WPI; 1988-049866/07.
 N-PSDB; AAN80446.

New DNA sequences encoding modified factor VIII:C - with deletion of DNA encoding maturation polypeptide, useful for high yield transformation.

Claim 3; Page 57-58-59-60; 97pp; English.

The entire sequence encoding the maturation polypeptide of factor VIII:C is deleted, i.e. Arg 740-Glu 1449.

The full length factor VIII:C cDNA has two changes at residue 1880 C to G and at residue 1881 C to G. The C to G change at amino acid residue 1880 C to G (Leu 242 and C to G) change at amino acid residue 1880 (Phe to Leu).

The product is produced in approx. 20 times higher yields than previous recombinant produced factor VIII:C and are easily purified. The peptide is used for treating haemophilia A, both C and CC.

See also AAN80444 and AAN80447.

(Updated on 25-MAR-2003 to correct PA field.)

Sequence 1425 AA;

Query Match 98.6%; Score 7581; DB 9; Length 1425;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 1421; Conservative 0; Mismatches 3; Indels 14; Gaps 1;

QY 1 ATRYTGLVGLSDHMGSLGELVQDAEPFPPKSPFNTSVVYKTLFVETVHLFN 60
 2 ATRYTGLVGLSDHMGSLGELVQDAEPFPPKSPFNTSVVYKTLFVETVHLFN 61

QY 61 TAKPFPWMLGLFTQAEVDTVVITLKNWASHVSLVAVGVYKASGEAYDDQSQ 120
 62 TAKPFPWMLGLFTQAEVDTVVITLKNWASHVSLVAVGVYKASGEAYDDQSQ 121

QY 121 REKEDKVPQSGSHYTWQLEKNGSPASDPLZATYSLSHVQVLYQNSGLGALLVC 180
 122 REKEDKVPQSGSHYTWQLEKNGSPASDPLZATYSLSHVQVLYQNSGLGALLVC 181

QY 181 EQLAEVETQTLHETFLFVAPDEKSGHSTQNTWQDQASQANWQVQVNR 240
 182 EQLAEVETQTLHETFLFVAPDEKSGHSTQNTWQDQASQANWQVQVNR 241

QY 241 EQLQGLCHRSKSVQVYVQWQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 300
 242 EQLQGLCHRSKSVQVYVQWQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 301

QY 301 WYQDPLQCHRSKSVQVYVQWQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 360
 302 WYQDPLQCHRSKSVQVYVQWQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 361

QY 361 DNNRSGPQIQRWQV 420
 362 DNNRSGPQIQRWQV 421

QY 421 RYKRVPMATDITNTREAIQRESGILGLLQVGVDTLLIFPQDASRPNTIYFHI 480
 422 RYKRVPMATDITNTREAIQRESGILGLLQVGVDTLLIFPQDASRPNTIYFHI 481

QY 481 TDVREYSSRLPQV 540
 482 TDVREYSSRLPQV 541

QY 541 RDAASGLGLPLCTYESVDQNGQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 600
 542 RDAASGLGLPLCTYESVDQNGQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 601

QY 601 VQLEDEPQASHMSEINGVDFSLQVSLVCLHEVAVYVWYLSGAQDPLSVFSSYTFKH 660
 602 VQLEDEPQASHMSEINGVDFSLQVSLVCLHEVAVYVWYLSGAQDPLSVFSSYTFKH 661

QY 661 NVYEDTLTPFSSGTVMKSNPLGLQCHNSPQVQVQVQVQVQVQVQVQVQVQVQV 720

```

Db 662 KNYTBDTLFFPGSTVPMHNPGLMILGCHNSDFPBGOWTALLKVSCKRKTQDYE 721
Qy 721 DSEYDIAILLKNNALIFPSSQNPVKKHQRKIRITLQDDEIDYDTTISVENKK 780
Db 722 USHDIASILLKNNALP-----RITLTLQSDDEIDYDTTISVENKK 767
Qy 781 EFDYDYEDENSPSPFOKTHYFFIAVERLMDYGMSSSHVLRNRAQSGSVPOFKVV 840
Db 788 EFDYDYEDENSPSPFOKTHYFFIAVERLMDYGMSSSHVLRNRAQSGSVPOFKVV 827
Qy 841 FOSEFDGSDTQPLVGELENHGLGQPYRAEVEDNIMWTFPMQASRPFYSLSISYEB 900
Db 828 FOSEFDGSDTQPLVGELENHGLGQPYRAEVEDNIMWTFPMQASRPFYSLSISYEB 887
Qy 901 DORQAGPRKIFVNEKTYFVQVHMAPTKDEFCQWAFPSVDLEKDVHSLGP 960
Db 888 DORQAGPRKIFVNEKTYFVQVHMAPTKDEFCQWAFPSVDLEKDVHSLGP 947
Qy 961 LLYCHTITLNAHGRQVTOVEFALPTTFQETKSMYTFENMENCASCHIMEDTFKE 1020
Db 948 LLYCHTITLNAHGRQVTOVEFALPTTFQETKSMYTFENMENCASCHIMEDTFKE 1007
Qy 1021 NYRPHATNYITQTLFGLWMAOODIRKXVLLNGENENIHSHPGHTVTKKRYNA 1080
Db 1008 NYRPHATNYITQTLFGLWMAOODIRKXVLLNGENENIHSHPGHTVTKKRYNA 1067
Qy 1081 LYNLYPQVFEVEMLSKAGTHRYECLIGELHAGMSTLEUVANWCOTRQMAQSHID 1140
Db 1068 LYNLYPQVFEVEMLSKAGTHRYECLIGELHAGMSTLEUVANWCOTRQMAQSHID 1127
Qy 1141 FOITASQVQNAKPLARLVKSGSINAMSTKEPFSMKVLLAPMIIHGIKTQRKQFS 1200
Db 1128 FOITASQVQNAKPLARLVKSGSINAMSTKEPFSMKVLLAPMIIHGIKTQRKQFS 1187
Qy 1201 SLVISOITWISLOKXWYTGNSGTGLAVTFGVDSGSKHNIENPPIIARYILRPT 1260
Db 1188 SLVISOITWISLOKXWYTGNSGTGLAVTFGVDSGSKHNIENPPIIARYILRPT 1247
Qy 1261 HYSISTLMEILAGDUNSCMPLGHSKALISDAQITASSTFPMFATNGPSKARHQAQ 1320
Db 1248 HYSISTLMEILAGDUNSCMPLGHSKALISDAQITASSTFPMFATNGPSKARHQAQ 1307
Qy 1321 RSNMWRQNNPKNLQVDFQKMYGTQOQVKSLLSNVYKEFTLSSSQDQOMTLP 1380
Db 1308 RSNMWRQNNPKNLQVDFQKMYGTQOQVKSLLSNVYKEFTLSSSQDQOMTLP 1367
Qy 1381 FONGKVVQFQNDGSPFVNSLDPLIARYILHIFQSWHQIALRMEVTLGCEADY 1438
Db 1368 FONGKVVQFQNDGSPFVNSLDPLIARYILHIFQSWHQIALRMEVTLGCEADY 1425

```

Search completed: December 9, 2003, 17:02:12
 Job time : 69 secs

Db 301 MDLQGLFLFCHISHSQIGQGEAYVYKVDQCEPQLAKMKNBEADYDDUTLDSBROVYRF 360
 QY 361 DDNDSGFSQIRVADKQSPKTVWYHTAAEEDWDYPLADPDRESKYSQYLANGPQIG 420
 Db 361 DDNDSGFSQIRVADKQSPKTVWYHTAAEEDWDYPLADPDRESKYSQYLANGPQIG 420
 QY 421 RYKYKAFWATDDETFREAAQESGELGELYGEVGTLLIFPNQAQSRPNYPHGI 480
 Db 421 RYKYKAFWATDDETFREAAQESGELGELYGEVGTLLIFPNQAQSRPNYPHGI 480
 QY 481 TVDRPLSRELPQGVKHLGLFPLGELFPGYVTVVYVGGTSPKDCPLKYSYSYVME 540
 Db 481 TVDRPLSRELPQGVKHLGLFPLGELFPGYVTVVYVGGTSPKDCPLKYSYSYVME 540
 QY 541 RLASGLGELPCLLYKESVQSGRQNDMSRWVLLFSYVDENESKYLTEINQPLNAG 600
 Db 541 RLASGLGELPCLLYKESVQSGRQNDMSRWVLLFSYVDENESKYLTEINQPLNAG 600
 QY 601 VLEDEPFGASINSGYVDSGLQVCLSVHAYVTLGACQDPLSVSFGTPEH 660
 Db 601 VLEDEPFGASINSGYVDSGLQVCLSVHAYVTLGACQDPLSVSFGTPEH 660
 QY 661 KNYVEEDTLPLPSCETVNSNGPGLLQNGSDFNQNGTALLKASCQKQDYVE 720
 Db 661 KNYVEEDTLPLPSCETVNSNGPGLLQNGSDFNQNGTALLKASCQKQDYVE 720
 QY 721 KNYVEEDTLPLPSCETVNSNGPGLLQNGSDFNQNGTALLKASCQKQDYVE 780
 Db 721 KNYVEEDTLPLPSCETVNSNGPGLLQNGSDFNQNGTALLKASCQKQDYVE 780
 QY 781 DSYEDISALSSNNALPERSQNPPLAKRQREITKTLAQDQEEIDYDDTISYEMXK 840
 Db 781 DSYEDISALSSNNALPERSQNPPLAKRQREITKTLAQDQEEIDYDDTISYEMXK 840
 QY 841 EDPLDYDDENGSPRESQKTRHTFTANERLMDYGNSSHLNRRQAGSEVQTKXV 840
 Db 841 EDPLDYDDENGSPRESQKTRHTFTANERLMDYGNSSHLNRRQAGSEVQTKXV 840
 QY 841 FQSTDGSTFQPLHGLNHLGLQPLTAEVDEINMTLTPQKGRPYSTSLISYE 900
 Db 841 FQSTDGSTFQPLHGLNHLGLQPLTAEVDEINMTLTPQKGRPYSTSLISYE 900
 QY 901 DQKQAGPQFVYDNETKTKYKQYHRAPLQDEFCQKATYSVDVLQKVSGLIG 960
 Db 901 DQKQAGPQFVYDNETKTKYKQYHRAPLQDEFCQKATYSVDVLQKVSGLIG 960
 QY 961 LVCCHTINLPAKQGVTVQEPALFTTIDETKSMYTFMERNCRAPCHIQDPTFE 1020
 Db 961 LVCCHTINLPAKQGVTVQEPALFTTIDETKSMYTFMERNCRAPCHIQDPTFE 1020
 QY 1021 NYRPAHNLQPLVQVADQDQRTYLLGNGSNNHIFSGVTVKAKRYEA 1080
 Db 1021 NYRPAHNLQPLVQVADQDQRTYLLGNGSNNHIFSGVTVKAKRYEA 1080
 QY 1081 LKYNYPGVETVEMLSKAGIWRVCLLGHLAGHSTLEFYVSKMCTPLGNASHED 1140
 Db 1081 LKYNYPGVETVEMLSKAGIWRVCLLGHLAGHSTLEFYVSKMCTPLGNASHED 1140
 QY 1141 FQSTQGVQGNAPLHWSGSIINAMSTKEPFSKIKVDLLAMIHGKTCQAQKFS 1200
 Db 1141 FQSTQGVQGNAPLHWSGSIINAMSTKEPFSKIKVDLLAMIHGKTCQAQKFS 1200
 QY 1201 SLVYSQITLWLSLQCKQVTKGNSGTGLPFFGVQSGSGLKRNPNPIIARILPHT 1260
 Db 1201 SLVYSQITLWLSLQCKQVTKGNSGTGLPFFGVQSGSGLKRNPNPIIARILPHT 1260
 QY 1261 HYSIESTLHMLMGCDLNSCPLNGMESKASDAQTASSTFTNFWATWSPKARILHJQ 1320
 Db 1261 HYSIESTLHMLMGCDLNSCPLNGMESKASDAQTASSTFTNFWATWSPKARILHJQ 1320
 QY 1321 RSNANRPOVNHKMLQVDFQKTKVQVTVQGVKSLLSKYSWYKEFLISSQDQKWTLP 1380
 Db 1321 RSNANRPOVNHKMLQVDFQKTKVQVTVQGVKSLLSKYSWYKEFLISSQDQKWTLP 1380
 QY 1381 FQNGKQVQSGPFPVNSQDPLKTEITLHPSQKQWQWQWQWQWQWQWQWQWQW 1438
 Db 1381 FQNGKQVQSGPFPVNSQDPLKTEITLHPSQKQWQWQWQWQWQWQWQWQWQW 1438

RESULT 2

US-08-683-8398-3
 / Sequence 3, 5
 / Application US/086838398
 / General Information:
 / APPLICANT: Ill. Charles . R. et al.
 / TITLE OF INVENTION: Use of Viral Cis-Acting Post-Transcriptional
 / Control Elements in the Regulation of Gene Expression of
 / TITLE OF INVENTION: Intracellular Genes Containing Near-Consensus Splice Sites
 / NUMBER OF SEQUENCES: 18
 / CORRESPONDENCE ADDRESS:
 / ADDRESS: 60 State Street, Suite 510
 / CITY: Boston
 / STATE: Massachusetts
 / COUNTRY: USA
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / SOFTWARE: Patent Release #1.0, Version #1.25
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/683,8398
 / FILING DATE: 11-MARCH-1996
 / CLASSIFICATION DATA:
 / APPLICATION NUMBER:
 / FILING DATE: INFORMATION:
 / NAME: Remillard, Jane E.
 / REGISTRATION NUMBER: 38,872
 / REFERENCE/DOCKET NUMBER: TTI-138
 / TECHNICAL CATION INFORMATION:
 / TELEPHONE: (617)227-7476
 / TELEFAX: (617)227-5941
 / INFORMATION FOR SEQ ID NO: 3:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 1438
 / TYPE: amino acid
 / TOPOLOGY: linear
 / MOLECULE TYPE: protein
 / US-08-683-8398-3
 Query Match 99.8%; Score 7674; DB 1; Length 1471;
 Best Local Similarity 99.0%; Pred. No. 0;
 Matches 1438; Conservative 0; Mismatches 0; Indels 14; Gaps 1;
 QY 1 ATATYTGAGVLSMDWDSGLGELPVDARFPKVPSPFPFNTSVYTKTLVFTVHLFN 60
 Db 20 ATATYTGAGVLSMDWDSGLGELPVDARFPKVPSPFPFNTSVYTKTLVFTVHLFN 79
 QY 61 TAPKPPFWGLGPTQAEYDVTVTTLKNWASHVHVAVGSTKASGAEYDDQSQ 120
 Db 80 TAPKPPFWGLGPTQAEYDVTVTTLKNWASHVHVAVGSTKASGAEYDDQSQ 139
 QY 121 EKKEDVDVPGSSHTVQVLENGPMSDPLCLTYSYLSHVDLVNLSGLGALLVC 180
 Db 140 EKKEDVDVPGSSHTVQVLENGPMSDPLCLTYSYLSHVDLVNLSGLGALLVC 199
 QY 181 ESKDAKENTQTLHVTFLVPRDKSKSHSTNSLAQDRDAASRPAQPTVGYNR 240
 Db 200 ESKDAKENTQTLHVTFLVPRDKSKSHSTNSLAQDRDAASRPAQPTVGYNR 259
 QY 241 ESKTCCKKAVPQVYVGNQTRKAVSTFLKSHPPVVRKRSKASLESTPLKATLL 300
 Db 260 ESKTCCKKAVPQVYVGNQTRKAVSTFLKSHPPVVRKRSKASLESTPLKATLL 319
 QY 301 MDLQGLFLFCHISHSQIGQGEAYVYKVDQCEPQLAKMKNBEADYDDUTLDSBROVYRF 360
 Db 320 MDLQGLFLFCHISHSQIGQGEAYVYKVDQCEPQLAKMKNBEADYDDUTLDSBROVYRF 379

QY 361 DDNSPSFQIRVAAGKPKVHYIAABEDMDYAPVLAPDDSYSTQYLNNGPQIG 420
 DB 380 DDNGSFGIRSVAGKPKVHYIAABEDMDYAPVLAPDDSYSTQYLNNGPQIG 439
 QY 421 RYKAYRPMATYETFKRETAQESGILAPLGEVGTQILLIPKQASRPYNIPIGI 480
 DB 440 RYKAYRPMATYETFKRETAQESGILAPLGEVGTQILLIPKQASRPYNIPIGI 499
 QY 481 TDVRLPSRLPQVGHKGLDPPILGLCEIFKMYVTIVDEGPTSPDCLVRYYSFVME 540
 DB 500 TDVRLPSRLPQVGHKGLDPPILGLCEIFKMYVTIVDEGPTSPDCLVRYYSFVME 559
 QY 541 TDRLASGLIGPLLCYKESVQGRQNDSDNWLFSVDENESYLTENIQEPLNPMAG 600
 DB 560 TDRLASGLIGPLLCYKESVQGRQNDSDNWLFSVDENESYLTENIQEPLNPMAG 619
 QY 601 VQEDPEFQASNTWESNGYFQSLQALCSLHVAWYLLSICQDTELVSFSGSTGHI 660
 DB 620 VQEDPEFQASNTWESNGYFQSLQALCSLHVAWYLLSICQDTELVSFSGSTGHI 679
 QY 661 KMYEDPTLPPSGPTVPMGNSQVGLMLGQNSQFMDKALPKSACQNTQVYIE 720
 DB 680 KMYEDPTLPPSGPTVPMGNSQVGLMLGQNSQFMDKALPKSACQNTQVYIE 739
 QY 721 PVEDVITLGLVILASQNNALPERSQSON -----PVEKNSQVITLGLQSGE 766
 DB 740 PVEDVITLGLVILASQNNALPERSQSONPSTQCFQNFATPVEKNSQVITLGLQSGE 799
 QY 767 EIVDQITLVNPKHQRVYDDEBNQSRTQKNTMYTIANVELMVOHSSPSWLN 826
 DB 800 EIVDQITLVNPKHQRVYDDEBNQSRTQKNTMYTIANVELMVOHSSPSWLN 859
 QY 827 RQSGSVDFKVPQDFDQSFQPLATGELNEHLGLGSLTAAYVEDNIMVTFNQA 886
 DB 860 RQSGSVDFKVPQDFDQSFQPLATGELNEHLGLGSLTAAYVEDNIMVTFNQA 919
 QY 887 RPYFSYSLIETEDQGAEPNTPVKNETKITFKYQORHAPTKQDFCKANVPSD 946
 DB 920 RPYFSYSLIETEDQGAEPNTPVKNETKITFKYQORHAPTKQDFCKANVPSD 979
 QY 947 VQLEQVHSGILQILCHVNTLAPAGQVTVQEFALFTPIDETKSYHTFEMNERCK 1006
 DB 980 VQLEQVHSGILQILCHVNTLAPAGQVTVQEFALFTPIDETKSYHTFEMNERCK 1039
 QY 1007 APCNTQMEDPFEKNTFRPAINGITMDGLVMAQOQRHWTLSGNSNHSIHPSG 1066
 DB 1040 APCNTQMEDPFEKNTFRPAINGITMDGLVMAQOQRHWTLSGNSNHSIHPSG 1099
 QY 1067 HPTVTEGSEYQALYNTFCVPTVEKLRSAGTNRVCLLGEHLHAGSTPLVYSNK 1126
 DB 1100 HPTVTEGSEYQALYNTFCVPTVEKLRSAGTNRVCLLGEHLHAGSTPLVYSNK 1159
 QY 1127 COTPLGASGHIRDFQITASGQVQOAPKLARLHSGSTNSMSTKPSFWIKDLAPMI 1186
 DB 1160 COTPLGASGHIRDFQITASGQVQOAPKLARLHSGSTNSMSTKPSFWIKDLAPMI 1219
 QY 1187 HIGITQGAOKHFSLSYISQITVMSLDKQKWOTGSGTNTLAPVGNVDSQSKNFI 1246
 DB 1220 HIGITQGAOKHFSLSYISQITVMSLDKQKWOTGSGTNTLAPVGNVDSQSKNFI 1279
 QY 1247 NPPIYATYLRPHYSIRSTLRMEJMGCLNSQWJQESKALSDQATYASSYTNMF 1306
 DB 1280 NPPIYATYLRPHYSIRSTLRMEJMGCLNSQWJQESKALSDQATYASSYTNMF 1339
 QY 1307 ATWSPSARHLLQGEENAPQVQVNTKMLQDQNTQVQVQKGLTQVQKGLTQVQK 1366
 DB 1340 ATWSPSARHLLQGEENAPQVQVNTKMLQDQNTQVQVQKGLTQVQKGLTQVQK 1399
 QY 1367 LIESSQDQKQNTLFPQNGQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 1426
 DB 1400 LIESSQDQKQNTLFPQNGQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 1459

QY 1427 MEVLACBQADLY 1438
 DB 1460 MEVLACBQADLY 1471

RESULT 3

US-08-882-083-2
 ; Sequence 2, Application US/0882083
 ; Patent No. 5,852,852
 ; GENERAL INFORMATION:
 ; APPLICANT: VOORBERG, Johannes J.
 ; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
 ; NUMBER OF SEQUENCES: 17
 ; COUNTRY: Netherlands
 ; ADDRESS: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; MEDIUM SIZE: 5.25 inch
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: 08/082,083
 ; FILING DATE: 08/08/82,083
 ; CLASSIFICATION: 514
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/558,107
 ; FILING DATE: 08/08/82,083
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ISACSON, John P.
 ; REGISTRATION NUMBER: 33,715
 ; ADDRESS: 1000 K Street, N.W., Suite 500
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 672-5300
 ; TELEFAX: (202) 672-5399
 ; FLEXION: 904138, ID NO. 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1661 amino acids
 ; TYPE: amino acid
 ; MOLECULAR TYPE: protein
 ; US-08-882-083-2

Query Match 96.71, Score 9134, DE 2; Length 1661;
 Best Local Similarity 86.04, Pos. No. 0,
 Matches 1418; Conservative 4; Mismatches 10; Indels 216; Gaps 4;

QY 1 ATNRVQGNVLSQNTMSQSGELVQAPFPFVSPFNTSVTACTLPVFTVHLN 60
 DB 20 ATNRVQGNVLSQNTMSQSGELVQAPFPFVSPFNTSVTACTLPVFTVHLN 79
 QY 61 TACPFPFMGLGPTTQAEVVDVTILKQWASHPSLVHAGVSTKASGAYDDQSQ 120
 DB 80 TACPFPFMGLGPTTQAEVVDVTILKQWASHPSLVHAGVSTKASGAYDDQSQ 139
 QY 121 RECDQNVFGSGSTTVMQIKENGPMASDPLCTITSLSHVDVQNSGLIGALLVC 180
 DB 140 RECDQNVFGSGSTTVMQIKENGPMASDPLCTITSLSHVDVQNSGLIGALLVC 199
 QY 181 EGSLAKETQTLKFLFLFAPVDSKSHSTNSLQKQDAQANAPQPTVNGVNR 240
 DB 200 EGSLAKETQTLKFLFLFAPVDSKSHSTNSLQKQDAQANAPQPTVNGVNR 259
 QY 241 SLQGLAGCHREYVTHVINGTPTVHVSIFLECHTFLVNRHQASLEISPTFLAQTLL 300
 DB 260 SLQGLAGCHREYVTHVINGTPTVHVSIFLECHTFLVNRHQASLEISPTFLAQTLL 319
 QY 301 NDLQGLFLCHSSHQDQKQWYKQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 360

140 REKEDKVFQSGSHYTVQVLKNGMAQSPCLITYSLSHVVLVQNLNGLIGALLVCH 199
 181 EGSLAKNKTOTLWELLTFLVDPVCKNSHSTFQGLHQRDASASNAWPMWYKVR 240
 200 EGSLAKNKTOTLANKFLVFLVDFGSLMSHSTFNSLAQRDASASNAWPMWYKVR 259
 241 SLKSLFCHRSYKWKVYKMGVYKFLASFLASGHTFLVNRHSLASLSPITVPAQLL 300
 260 SLPLGLTCHRSYKWKVYKMGVYKFLASFLASGHTFLVNRHSLASLSPITVPAQLL 319
 301 MLGQFLFCHLSSHSDHOMENAYKVCSPFSEPLEMNRHSDYDUDLTSMSVVPF 360
 320 MLGQFLFCHLSSHSDHOMENAYKVCSPFSEPLEMNRHSDYDUDLTSMSVVPF 379
 361 IDNDSSEFQTSVAKGKPTWYHVAABEDMDATFLVLPADRSYKSTGYLANKGQRTG 420
 380 IDNDSSEFQTSVAKGKPTWYHVAABEDMDATFLVLPADRSYKSTGYLANKGQRTG 439
 421 KYKQKPMATYDTEFKREAQHEGSLGPELLVCEGDTALLFKQAGRYTHYPHGI 480
 440 KYKQKPMATYDTEFKREAQHEGSLGPELLVCEGDTALLFKQAGRYTHYPHGI 499
 481 TVDRLVLSRLPQGVGKLPFLPGSTFKYKTVYVDEGFTSDPCLTRYYSVPMSE 540
 500 TVDRLVLSRLPQGVGKLPFLPGSTFKYKTVYVDEGFTSDPCLTRYYSVPMSE 559
 541 XOLAAGLITGLPYCYEVSQVQRIHMDGNVILFVDFDRNSGVLITENIQRFLNAPG 600
 560 XOLAAGLITGLPYCYEVSQVQRIHMDGNVILFVDFDRNSGVLITENIQRFLNAPG 619
 601 VOLEDPFQASNPRIHNGVYFVPSLQSLVCLSEVAVATYLSIGQDTFLSVFSGYTFH 660
 620 VOLEDPFQASNPRIHNGVYFVPSLQSLVCLSEVAVATYLSIGQDTFLSVFSGYTFH 679
 661 KMYEDTLLPFPSSEIVSMENPCVLGCHNSDFNRGHTALLKVSCK---DKNTGD 717
 680 KMYEDTLLPFPSSEIVSMENPCVLGCHNSDFNRGHTALLKVSCKTFGEEDDD 739
 718 YV-----EUSYEDISANLYSKNAIETRSF----- 742
 740 YDLSKLTFSBDDYDIDI-----VDSLEBRSFQSNRHPSTKQKQFVATTIPENDIKET 793
 743 ----- 742
 794 DFWFAHTPMPIKQVSSDOLLAKULASPTPHGLSLQSLQAKVETFSDDSPQASDNN 853
 743 ----- 742
 854 SLSEWTFHPLHSGHGWMTFPEGLQLRLEKLGTTADJPLAWNHVYGTQIPKEWSQE 913
 743 ----- 742
 914 ISPEKTAFKKDTOTLSLACENSLAALINEGQNPRIEVTWAKQRTSELCSNPVLK 973
 743 ----- 742
 975 KQRTWTFHTLQSGEIVTWTPTLSPKAWKEDFTVYVNRHNGFSGFQKTHYPIAVE 810
 974 KQRTWTFHTLQSGEIVTWTPTLSPKAWKEDFTVYVNRHNGFSGFQKTHYPIAVE 1033
 811 RLYDQNGSSFTLNRAGSGVTPQKVVTFQSTFDSPTQPLTGLNHLHGLGPTVR 870
 1034 RLYDQNGSSFTLNRAGSGVTPQKVVTFQSTFDSPTQPLTGLNHLHGLGPTVR 1093
 871 APVENDVMTFTRQAGVTFYSVSLTSTREDDQALSPRQNPVETKTYPKVQVHQA 930
 1094 APVENDVMTFTRQAGVTFYSVSLTSTREDDQALSPRQNPVETKTYPKVQVHQA 1153
 931 PTQSPFCANATFSDVLEKQVHGLIGLAVCHNTLNPARGQVTVQEPALPTFFD 990
 1154 PTQSPFCANATFSDVLEKQVHGLIGLAVCHNTLNPARGQVTVQEPALPTFFD 1213
 991 TSWTSTFNNKRCACPICHMDPTFKFENYFALNGVIMDTLPLVMAQQRIRWTL 1050
 1214 TSWTSTFNNKRCACPICHMDPTFKFENYFALNGVIMDTLPLVMAQQRIRWTL 1273

RESULT 5

US-09-243-539-2

; Sequence 2, Application US/09243539
 ; Patent No. 6130253
 ; Inventor: J. COHEN;
 ; APPLICANT: VOCBERG, Johannes J.
 ; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESS: 3000 K Street, N.W., Suite 500

CITY: Washington

STATE: D.C.

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/243,539

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/558,107

ATTORNEY/AGENT INFORMATION:

NAME: ISAACSON, Joan F.

REGISTRATION NUMBER: 33,715

REFERENCE/JOINT NUMBER: 30472/212

TELEPHONE: (202)672-5300

TELEFAX: (202)672-5399

TELEX: 904136

INFORMATION TO NO. 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1661 amino acids

TYPE: amino acid

MOLECULE TYPE: Linear

US-09-243-539-2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 404-815-6508

TELEFAX: 404-815-6555

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 2332 amino acids

TYPE: amino acid

STANDARDNESS: single

COMPLETENESS: full length

MOLECULE TYPE: protein

HYPOTHEICAL: NO

ANTI-SENSE: NO

TRANSLATION: N-terminal

ORIGINAL SOURCE: Homo sapien

ORGANISM: Homo sapien

TISSUE TYPE: Liver cDNA sequence

US-07-864-004B-4

Query Match 94.13; Score 7234; DB 1; Length 2332;

Best Local Similarity 61.71; Pred. No. 0;

Matches 1438; Conservative 0; Mismatches 0; Indels 894; Gaps 1;

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

1 ATNRYLGLAVLSNDWMSDGLGELPVDARPPRVKFSPTVYVYKTLFVETVHLN 60

Qy	743	-----	742
Db	901	TSSLPPSPWVHQSDQDTTL-FQKXSKSPJTESQPLSEJBEENNUSKJLESQJANQJSSW	960
Qy	743	-----	742
Db	961	QKMSVSTESGRLFKGDHAGBALITKDNALFKWSLSLLTWTKNNSNANWKTHTDQSL	1020
Qy	743	-----	742
Db	1021	LINSSTPKWJLNSSTFKYVPLHDPALADNATALAHMSNMTSTSSNNWQOK	1080
Qy	743	-----	742
Db	1081	KEBPIPPDQCNPMKSPFKMLIPFSARWIOPTHKNSQXGSPKQJVSJGFEKSVG	1140
Qy	743	-----	742
Db	1141	QNFLESBKVVYVGGEFTDQVLGHEKMPFSSNNFLNLDLHNNTNIBKIQBIEK	1200
Qy	743	-----	742
Db	1201	KETLQENVVFLQIHTVOTNPKNPLASTHQNVGSEGAVALPQPSLNDSTNR	1260
Qy	743	-----	742
Db	1261	TKKTHAFHSXGEBENLEGJNQTKVEKIACTRI-SPTNSQNVTOBSKALQFEL	1320
Qy	743	-----	742
Db	1321	PLBETELKXILVNDTQWNSKNMHLFTSTQIDTNEKKAITQPSJLCLTBSHSI	1380
Qy	743	-----	742
Db	1381	PQARSPLFIAKVSJFSPSIRHYLTRVLFQDNSSHLPAASYKQSGVSHFLQACX	1440
Qy	743	-----	742
Db	1441	NULSIALTLWTGDBQSVGSLQTSNTSVTKYKNTVLPKPLDNTSGVLLPKVHI	1500
Qy	743	-----	742
Db	1501	YKQDLPTSTXSGSPHLLVREGSLQGTSEALKNEANRPQKVPJLVATSSANTPSK	1560
Qy	743	-----	742
Db	1561	LIDPLANVGTQIPKEMXGKSPKTKFKUOTLISJACSENHJAJANQGNKP	1620
Qy	743	-----	742
Db	1621	ELEVWANGORTELCQNPQJLQDQOEDITLSDQDEIDTDDTISYMKKEDDIY	1680
Qy	787	DEBNQSPSPKCHTRYIAAVERLWYQWMSSSHPHLABRAGQSVDFPKVWVFQD	846
Db	1681	DEBNQSPFOKCHTRYIAAVERLWYQWMSSSHPHLABRAGQSVDFPKVWVFQD	1740
Qy	847	GSFTQPLFRSELNHLGLQPIABVDENLWTFPQKQSPFVPSLSLISYENQDCA	906
Qy	743	-----	742
Db	907	GSFTQPLFRSELNHLGLQPIABVDENLWTFPQKQSPFVPSLSLISYENQDCA	1800
Qy	961	EPKNNVPLNENKTIEMKYQCHMAPKQDPCAKAWSDVDLCKQVSHSLQPLAVCH	966
Db	1801	EPKNNVPLNENKTIEMKYQCHMAPKQDPCAKAWSDVDLCKQVSHSLQPLAVCH	1860
Qy	967	NLTPAQHROVTOVEBAPLFTIPDTSKNTPTENNEBCASCHTOMQDPTFVNSRHA	1026
Qy	1027	INGVIMPTLEGLWQADQRETKLQNSNTNNSHUSGHWKTRKSNQANWJY	1086
Db	1021	INGVIMPTLEGLWQADQRETKLQNSNTNNSHUSGHWKTRKSNQANWJY	1086

RESULT 8
US-08-212-133A-2
; Sequence 2, Application US/08212133A
; Patent No. 5663060

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
SOFTWARE: PATENTIN Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
SUBJECT: 108/212.133A
PILING NUMBER: 135, 1394
PILING DATE: 07-APR-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
SUBJECT: 108/964,004
PILING NUMBER: 135, 1394
PILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Babot, Patricia L.
FIRM: 284
REFERENCE/DOCKET NUMBER: 76677
TELECOMMUNICATION INFORMATION:
TELEPHONE: 604-572-6508
FAX: 604-572-6508
INFORMATION FOR SEQ ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
FEATURES:
STRANDINESS: single
TOPOLOGY: linear
N-TERMINAL TYPE: proein
C-TERMINAL TYPE: HIS
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORGANISM SOURCE: sapient
TISSUE TYPE: Liver cDNA sequence

```

US-08-212-133A-2

Query Match 94.1%; Score 7234; DB 1; Length 2332;
 Best Local Similarity 61.7%; Pred. No. 0;
 Matches 1439, Conservative 0; Mismatches 0; Indels 894; Gaps 1;

QY 1 ATRRYLGAVALSHNDYMDGELGELPDARPRVPKSPFPFTSVYKTLPLVFETVHLPM 60
 DB 1 ATRRYLGAVALSHNDYMDGELGELPDARPRVPKSPFPFTSVYKTLPLVFETVHLPM 60
 QY 61 TAKRPPMGLGLOFTQABVDTOTVITLAWASHPVSHVSNVKAASGASVADQDSO 120
 DB 61 TAKRPPMGLGLOFTQABVDTOTVITLAWASHPVSHVSNVKAASGASVADQDSO 120
 QY 121 REKKDKVFGGSHYVWQULKENGDMASOELCLTYSYSHWLYGALNSCLGALLVCE 180
 DB 121 REKKDKVFGGSHYVWQULKENGDMASOELCLTYSYSHWLYGALNSCLGALLVCE 180
 QY 181 EGSJAEKOTYTLHEKTLFAPVDFGKQWSEYFNGSLNODRANASAPVQKQVYVR 240
 DB 181 EGSJAEKOTYTLHEKTLFAPVDFGKQWSEYFNGSLNODRANASAPVQKQVYVR 240
 QY 241 ELQGLICHRKSVYVWYVQGTGTHVTHLSEHFTLVNRRKASLEISPTFVETPL 300
 DB 241 ELQGLICHRKSVYVWYVQGTGTHVTHLSEHFTLVNRRKASLEISPTFVETPL 300
 QY 301 YKLGSLFLICHSKSGSNVYVQVQKQWSEYFNGSLNODRANASAPVQKQVYVR 360
 DB 301 YKLGSLFLICHSKSGSNVYVQVQKQWSEYFNGSLNODRANASAPVQKQVYVR 360
 QY 361 DDMSSQTSIRSVNKKVYVWYVQGTGTHVTHLSEHFTLVNRRKASLEISPTFVETPL 420
 DB 361 DDMSSQTSIRSVNKKVYVWYVQGTGTHVTHLSEHFTLVNRRKASLEISPTFVETPL 420
 QY 421 RYKAVRPMATDTFTREALOESGLGLOPLGELVGTGTLILIPKQASRPVYTHCI 480
 DB 421 RYKAVRPMATDTFTREALOESGLGLOPLGELVGTGTLILIPKQASRPVYTHCI 480
 QY 481 TVRPLSRPLGVLKLOPFLPGELFKMTVTVEDGTFASDPRCLTYVSYVWME 540
 DB 481 TVRPLSRPLGVLKLOPFLPGELFKMTVTVEDGTFASDPRCLTYVSYVWME 540
 QY 541 RLASGLIGLICKESVQSGNQTMDKNNVILFVDFRNSMYITENIQRLEPNAG 600
 DB 541 RLASGLIGLICKESVQSGNQTMDKNNVILFVDFRNSMYITENIQRLEPNAG 600
 QY 601 VOLEDPQANLMSINGVYVQVQKQWSEYFNGSLNODRANASAPVQKQVYVR 660
 DB 601 VOLEDPQANLMSINGVYVQVQKQWSEYFNGSLNODRANASAPVQKQVYVR 660
 QY 661 KVVYBTLPLFFSGEYVFNMSMNPGLMILGCHNSDFRNRKALLKVSCKDKTYDE 720
 DB 661 KVVYBTLPLFFSGEYVFNMSMNPGLMILGCHNSDFRNRKALLKVSCKDKTYDE 720
 QY 721 DSYEDIATLSDNNALPERSPSONSRPSTKQCFNATIPENDIKTPMFAHRTMP 780
 DB 721 DSYEDIATLSDNNALPERSPSONSRPSTKQCFNATIPENDIKTPMFAHRTMP 780
 QY 743 ----- 742
 DB 743 ----- 742
 QY 781 KIQVSSDMLLAQSPFHGLSLDQKAVETFDSDPQADSNNSLGENTHFPQ 840
 DB 781 KIQVSSDMLLAQSPFHGLSLDQKAVETFDSDPQADSNNSLGENTHFPQ 840
 QY 743 ----- 742
 DB 743 ----- 742
 QY 841 LHHSGVMTFESGQJRLNEKLOTTAVELKXLOLKVFSSTNSLNLSITPSNLAAGTN 900
 DB 841 LHHSGVMTFESGQJRLNEKLOTTAVELKXLOLKVFSSTNSLNLSITPSNLAAGTN 900
 QY 743 ----- 742
 DB 743 ----- 742
 QY 901 TSSLPNPFVHDSQULTLTFOKSGSPSTESGQJSLSENNUSKLESLGANSSESN 960
 DB 901 TSSLPNPFVHDSQULTLTFOKSGSPSTESGQJSLSENNUSKLESLGANSSESN 960
 QY 743 ----- 742
 DB 743 ----- 742
 QY 961 GKNVSTESGLFKGKANGFALLTKDNLKFKVSISSLAKTKSNSATKXTHIDGSL 1020
 DB 961 GKNVSTESGLFKGKANGFALLTKDNLKFKVSISSLAKTKSNSATKXTHIDGSL 1020

QY 743 ----- 742
 DB 1021 LIHNSPSWQVILSUTEFKAVTEFLHDMKMKATALLHNSKNTSSQKRWQCK 1080
 QY 743 ----- 742
 DB 1081 KEGFTPPDAQFDMSEFOOLFLPESARWORTGHGRLNSQSPSPQOLVSLGPEKSVG 1140
 QY 743 ----- 742
 DB 1141 QVFLSEKNNVKGSEFTDQVGLKVPSEBNEFLJWLJNLHFNTHNCKKCOEIEK 1200
 QY 743 ----- 742
 DB 1201 KETTLQENVLPOLHVTGCTGPKWLLPSTLSTQNVSSYEGAVLQDFRSINDSTR 1260
 QY 743 ----- 742
 DB 1261 TQGTATFESKKEENLGLNGLOKQIVKCACTTCTSPNTSQNVPTQBSIKALQPL 1320
 QY 743 ----- 742
 DB 1321 PJEFTELEKRIIVDQTSQWNSKMKULTPSTLQIDWNKEKCAITQSPJLSDCLTSHSI 1380
 QY 743 ----- 742
 DB 1381 PQANRSLPJMKVSPFPPIPYTRVLPQNSHLSALPASFKQSGVQSSHFQGNK 1440
 QY 743 ----- 742
 DB 1441 NKSJLALTEMGQREVSUGTSNVTYKVTYVLPKPLPFSOKVLLPKVI 1500
 QY 743 ----- 742
 DB 1501 YQKULPFTSNGSPCHLDLVUGSLQTECAKKNENRPGKVPFLVATESAATPSK 1560
 QY 743 ----- 742
 DB 1561 LLOPLANDHNYTOQIFSEKSGQBSPEKAPKCKOTILSLNACSNHATAAINBQKRP 1620
 QY 743 ----- 742
 DB 1621 EKVYVAKQORTELCSNTPKAPKQORERTITLQSQBERIIDYDQTSVENKEDPDIY 1680
 QY 787 DEBQESPOKCTKHTVTAANVELWYDGMSSHWLJNRASQGVQKQVQKQVQKQV 846
 DB 1681 DEBQESPOKCTKHTVTAANVELWYDGMSSHWLJNRASQGVQKQVQKQVQKQV 846
 QY 847 GSVTOPVYRGLNETHLGLQVTRAEVDNLTWTFNQAASPFYSGSLSVBERQQA 906
 DB 1741 GSVTOPVYRGLNETHLGLQVTRAEVDNLTWTFNQAASPFYSGSLSVBERQQA 906
 QY 907 EBNKQVTEBNSNTVYKQENBAPKQKQSCANVSDVQKQVQKQVQKQVQKQV 966
 DB 1801 EBNKQVTEBNSNTVYKQENBAPKQKQSCANVSDVQKQVQKQVQKQVQKQV 966
 QY 967 YLANAIGQNTVDEFLPFTITDTSKATTFENMCAFCIONDPTPKNTAFPA 1020
 DB 1861 YLANAIGQNTVDEFLPFTITDTSKATTFENMCAFCIONDPTPKNTAFPA 1020
 QY 1027 INCYITLPLGVMADQRTWLLSGNENHSHFSGRUFTFKQKRYKALALYLP 1086
 DB 1921 INCYITLPLGVMADQRTWLLSGNENHSHFSGRUFTFKQKRYKALALYLP 1086
 QY 1087 GVFVEMLPSSQGTWVRECLTGBHJAGNCTLFWNSKCTPLAMASGHKDFQPTAS 1146
 DB 1981 GVFVEMLPSSQGTWVRECLTGBHJAGNCTLFWNSKCTPLAMASGHKDFQPTAS 1146
 QY 1147 QVQWAPKALALHYGSIINASTKEPFWIKYDLAPMLIHGKTCQAKQKFSLSIQ 1206
 DB 2041 QVQWAPKALALHYGSIINASTKEPFWIKYDLAPMLIHGKTCQAKQKFSLSIQ 2100

QY 1207 FIWMSLDKQKQWTRNGSTOTLWVFFGVDSICXKRNFPPIIAXIYRUPHTHSIRS 1266
 Db 2101 FIWMSLDKQKQWTRNGSTOTLWVFFGVDSICXKRNFPPIIAXIYRUPHTHSIRS 2160
 QY 1267 TURMELGDCNLSNCFMAGSKAISDAQITAGSFTFMNATSPSRLHLOQSNMR 1326
 Db 2161 TURMELGDCNLSNCFMAGSKAISDAQITAGSFTFMNATSPSRLHLOQSNMR 2220
 QY 1327 POWNMPKMLDQCFKQKTKVGVVGVVQVSLTSMYVATSLSSQDGHQTLFPQKRY 1386
 Db 2221 POWNMPKMLDQCFKQKTKVGVVGVVQVSLTSMYVATSLSSQDGHQTLFPQKRY 2280
 QY 1387 KYPQGNQDSFTFVNSDLPULRYLRHPOSNVQIALNMEVLAQEAQDY 1438
 Db 2281 KYPQGNQDSFTFVNSDLPULRYLRHPOSNVQIALNMEVLAQEAQDY 2332

RESULT 9

US-08-474-503-2

Sequence 2, Application US/08474503

Sequence 1, Application US/08474503

GENERAL INFORMATION:

APPLICANT: Emory University

TITLE OF INVENTION: Hybrid Human/Animal Factor VIII

INVENTOR: KALPATRICK & CODY

CORRESPONDENCE ADDRESS: 12

ADDRESS: 1100 Peachtree Street, Suite 2800

CITY: Atlanta

STATE: Georgia

COUNTRY: US

ZIP: 30309

COMPUTER READABLE FORM:

SEQUENCE TYPE: CDS

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

COMPLETION DATE: 08/474, 503

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

REFERENCE NUMBER: 29, 476

REGISTRATION NUMBER: EMU06CIR(3)

TELEPHONE: 404-815-5500

TELEFAX: 404-815-5500

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 232 amino acids

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FUNCTION: N-terminal

FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

TISSUE: Liver

US-08-474-503-2

Query Match 94.18; Score 7234; DB 1; Indels 894; Gaps 1;
 Matched 1438; Conservative 0; Mismatches 0;
 QY 1 ATATYTGAVLMDMAGSDGELVDNARPPVPKSPFTTSVYKTLFVHTVHLFN 60
 Db 1 ATATYTGAVLMDMAGSDGELVDNARPPVPKSPFTTSVYKTLFVHTVHLFN 60
 QY 61 TAKPFPWMLGLPTQAEVYVTVTLKQNASHPVSLAVGVSTWASGEAYDDQTSQ 120

Db 61 TAKPFPWMLGLPTQAEVYVTVTLKQNASHPVSLAVGVSTWASGEAYDDQTSQ 120
 QY 121 REKEDVYVPGSHHTYVWVLKNGMAMSDPCLITTSLSHVDVLYDNLGSLCALVCH 180
 Db 121 REKEDVYVPGSHHTYVWVLKNGMAMSDPCLITTSLSHVDVLYDNLGSLCALVCH 180
 QY 181 BSLAGKNTQTLKFLTAVPDPGSHSHSTKNSLQDRDAGAAAPPGHTTNYVNR 240
 Db 181 BSLAGKNTQTLKFLTAVPDPGSHSHSTKNSLQDRDAGAAAPPGHTTNYVNR 240
 QY 241 SLPLTLCGRKSTYVHVYVINGTTPVWVSIFLGGHTFVLRHSCASLETSITFTQCLL 300
 Db 241 SLPLTLCGRKSTYVHVYVINGTTPVWVSIFLGGHTFVLRHSCASLETSITFTQCLL 300
 QY 300 SLPLTLCGRKSTYVHVYVINGTTPVWVSIFLGGHTFVLRHSCASLETSITFTQCLL 300
 Db 300 SLPLTLCGRKSTYVHVYVINGTTPVWVSIFLGGHTFVLRHSCASLETSITFTQCLL 300
 QY 360 MDLQGFPLPCHSSHQCHMEAYKVYSCPEPLRMKNCEAYDDDLTSDSMQVRF 360
 Db 360 MDLQGFPLPCHSSHQCHMEAYKVYSCPEPLRMKNCEAYDDDLTSDSMQVRF 360
 QY 420 DDNPSFTQISRVACQKPTWVYTAABEDMDAPLAPDRSKYSLQYLNMPQRIG 420
 Db 420 DDNPSFTQISRVACQKPTWVYTAABEDMDAPLAPDRSKYSLQYLNMPQRIG 420
 QY 480 RKYKQVPMYVDTFTKTBRAQHSGLCLPLAYCGVDTLLTIPKQASRPYHPIGT 480
 Db 480 RKYKQVPMYVDTFTKTBRAQHSGLCLPLAYCGVDTLLTIPKQASRPYHPIGT 480
 QY 540 TDVPLSRPLKGVGLKADPLPGRI PKYKVTYVVGQTKSDQRCGLTVKSYSNVME 540
 Db 540 TDVPLSRPLKGVGLKADPLPGRI PKYKVTYVVGQTKSDQRCGLTVKSYSNVME 540
 QY 600 VLEDPFQNSIMHSINGVTVSLQLSVLHEVAVYLLSQATDPLAVFTSOTFPH 600
 Db 600 VLEDPFQNSIMHSINGVTVSLQLSVLHEVAVYLLSQATDPLAVFTSOTFPH 600
 QY 660 KMYEDTLTLPFSGSTVSNVPCMLVILGQNSDFNRGMLTKLVKSSCKQTYVTE 720
 Db 660 KMYEDTLTLPFSGSTVSNVPCMLVILGQNSDFNRGMLTKLVKSSCKQTYVTE 720
 QY 720 DSTDYDANLTKNNALPEPSF----- 742
 Db 720 DSTDYDANLTKNNALPEPSF----- 742
 QY 742 DSTDYDANLTKNNALPEPSF----- 742
 Db 742 DSTDYDANLTKNNALPEPSF----- 742
 QY 840 KLVNVSDDLMLLQSPFTHGLSLDQKATKTFDFSPGSLNSGLSENTHFRQ 840
 Db 840 KLVNVSDDLMLLQSPFTHGLSLDQKATKTFDFSPGSLNSGLSENTHFRQ 840
 QY 900 LHSGDMVTFPSGQLKMLKLGQTTATLKLQFVNSSTNNLSTIPSDNLAQTN 900
 Db 900 LHSGDMVTFPSGQLKMLKLGQTTATLKLQFVNSSTNNLSTIPSDNLAQTN 900
 QY 960 TSSGLQSPVHVYVDSGLTTLTKGKSSGFTLSQSGPLSLSENNDSKLSEGLAMQSSW 960
 Db 960 TSSGLQSPVHVYVDSGLTTLTKGKSSGFTLSQSGPLSLSENNDSKLSEGLAMQSSW 960
 QY 742 GRNVSTSGRLFKGGEAHPALLTKQNALFKVSIILKTKNTSNSTNRKTHIDGSL 742
 Db 742 GRNVSTSGRLFKGGEAHPALLTKQNALFKVSIILKTKNTSNSTNRKTHIDGSL 742
 QY 1080 LIENSPSVQVNLSDTFKFKYVPLMIDHMLMDKNTALNLRHNSKTYSSQNKENVOOK 1080
 Db 1080 LIENSPSVQVNLSDTFKFKYVPLMIDHMLMDKNTALNLRHNSKTYSSQNKENVOOK 1080
 QY 1140 KEGTTPDQNFQNSFPFQCLFIPESAKRTQHGNSLNSQGSQPSGLVAPKPSVVG 1140
 Db 1140 KEGTTPDQNFQNSFPFQCLFIPESAKRTQHGNSLNSQGSQPSGLVAPKPSVVG 1140
 QY 1200 QNLSEKRVVKGKGFQVGLKEMVFPSSMLTNLQNLHSHNTNQKIKQREIK 1200


```

RESULT 11
US-09-037-601-2
Sequence 2, Application US/09037601
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
INVENTOR: Lollar, John S.
CORRESPONDENCE ADDRESS: 40
STREET: 5370 Manhattan Circle Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM: disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PC-DOS/MS-DOS
PATENTIN RELEASE #1.0, Version #1.30
CLASSIFICATION: US/09/037,601
FILING DATE: 26-JUN-1996
CLASSIFICATION: INTL:
PRIORITY APPLICATION NUMBER: NO PCT/US94/13200
FILING DATE: 15-NOV-1994
PRIORITY APPLICATION NUMBER: US 08/212,113
FILING DATE: 07-APR-1992
PRIORITY APPLICATION NUMBER: US 07/864,004
APPLICANT NUMBER: 33,878
NAME: Lollar, John S.
NAME: Fechter, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 75-95F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/439-8089
TELEFAX: 303/439-8089
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
TYPE: amino acid
STRAND: single
TOPOLGY: not relevant
FUNCTION: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORGANISM: Homo sapiens
TISSUE TYPE: Liver
US-09-037-601-2
Query Match 94.13; Score 7234; DB 3; Length 3322;
Best Local Similarity 61.7%; Pred. No. 0;
Matches 1438; Conservative 0; Mismatches 894; Gaps 1;
QY 1 ATREYGVGLVSHDQMSDGLPVDSAPPRVPSKSPFVSQVKTIVYVWLVN 60
DB 1 ATREYVGLVSHDQMSDGLPVDSAPPRVPSKSPFVSQVKTIVYVWLVN 60
QY 61 TAPDPDPVGLGFTQAEVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTV 120
DB 61 TAPDPDPVGLGFTQAEVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTV 120
QY 121 REEDKDVYFGSHYVWVLENGWASDPLCLTISTASHVDLVNGLGIALVCR 180
DB 121 REEDKDVYFGSHYVWVLENGWASDPLCLTISTASHVDLVNGLGIALVCR 180
181 EGSJANKVTKLWKEFTLLFVVDGDSHSTSTNETVQDDAASASAPWPNVNR 240
DB 181 EGSJANKVTKLWKEFTLLFVVDGDSHSTSTNETVQDDAASASAPWPNVNR 240
241 ELPLQCHKSKSVTVVIGMTGTPDVSVSTFEGHTFLVNRHQASLEISPTFLAQTL 300
DB 241 ELPLQCHKSKSVTVVIGMTGTPDVSVSTFEGHTFLVNRHQASLEISPTFLAQTL 300
301 MDGLQFLPCHLSHSRCHGMAVYKVQSPCEPQLQMKNNSEADYDDDLTDSMVVRF 360
DB 301 MDGLQFLPCHLSHSRCHGMAVYKVQSPCEPQLQMKNNSEADYDDDLTDSMVVRF 360
361 DDNSPFIQIUSVACQVQVWHVTHAAZEDMDVAPVLAPDPSISQYVANGPQIG 420
DB 361 DDNSPFIQIUSVACQVQVWHVTHAAZEDMDVAPVLAPDPSISQYVANGPQIG 420
421 RYKAYRPMATVDTSTFTRAAIQHESGELGLVGEVGLTLIFNQASRPVNYPHGI 480
DB 421 RYKAYRPMATVDTSTFTRAAIQHESGELGLVGEVGLTLIFNQASRPVNYPHGI 480
481 TDVRLVSRRLKQVGHVADPFLPGRIKTKVTVYVGGTSDPRLQVRYSYFNWME 540
DB 481 TDVRLVSRRLKQVGHVADPFLPGRIKTKVTVYVGGTSDPRLQVRYSYFNWME 540
541 DLASGLGLPLATCYEEVDQGNQIMQDKNVLSFVPSDENRSHLTENQRTLPNAP 600
DB 541 DLASGLGLPLATCYEEVDQGNQIMQDKNVLSFVPSDENRSHLTENQRTLPNAP 600
601 VQLEDPFQASNIHBSINGVDSQSLSCVLRVAVWYLLSGQTDPLSVFSPSYTFH 660
DB 601 VQLEDPFQASNIHBSINGVDSQSLSCVLRVAVWYLLSGQTDPLSVFSPSYTFH 660
661 KAYVEDTLTLPSPSGTTFMKNENRCLATLQCHSDPDRNGMTALLKVSQCDKNGDYIE 720
DB 661 KAYVEDTLTLPSPSGTTFMKNENRCLATLQCHSDPDRNGMTALLKVSQCDKNGDYIE 720
721 DSYEDTSAYLLSKNNALIPRSEF 742
DB 721 DSYEDTSAYLLSKNNALIPRSEF 742
743 ----- 742
743 ----- 742
781 KIQVNSDGLMLLRQSPTHGLSLDQAEKTFSDPSFCAIDSNNSLSENTHFRQ 840
DB 781 KIQVNSDGLMLLRQSPTHGLSLDQAEKTFSDPSFCAIDSNNSLSENTHFRQ 840
841 LHSIDQNVTFPSGLQLANKELOTTAAVELKDLQKFSVSTSNLSLSTIFSDNLAQTDN 900
DB 841 LHSIDQNVTFPSGLQLANKELOTTAAVELKDLQKFSVSTSNLSLSTIFSDNLAQTDN 900
901 TSSGLPPSPNVHDSQLTTLFGKNSSPLETSOGCLSLSEENDSKLSEGLANQSSSW 960
DB 901 TSSGLPPSPNVHDSQLTTLFGKNSSPLETSOGCLSLSEENDSKLSEGLANQSSSW 960
961 GKVSSTESGELPKGKGHPALLATKONALFKVSIILAKTKNTSNNAKRWTHIDGSL 1020
DB 961 GKVSSTESGELPKGKGHPALLATKONALFKVSIILAKTKNTSNNAKRWTHIDGSL 1020
1021 LIENSFWNVNLLSTFEFKYVPLHDEMLADQKATALALHNSNKTISKKNWQVK 1080
DB 1021 LIENSFWNVNLLSTFEFKYVPLHDEMLADQKATALALHNSNKTISKKNWQVK 1080
1081 KEGPDPDQNPMSFPQQLFPEGSANWLTQTHQKONSINGOGHSPQLVGLGPRGVG 1140
DB 1081 KEGPDPDQNPMSFPQQLFPEGSANWLTQTHQKONSINGOGHSPQLVGLGPRGVG 1140
1141 QNPLSEKVVGVGGHFTVGLGKRVFSPSRNPLFTLNLUNLHNTHNQKIKQIEIK 1200
DB 1141 QNPLSEKVVGVGGHFTVGLGKRVFSPSRNPLFTLNLUNLHNTHNQKIKQIEIK 1200
1201 KEPLQENVLPQIHTVTKTNMKNLFLSTQKQVBSYEGAVPLQDFGLANDTVR 1260
DB 1201 KEPLQENVLPQIHTVTKTNMKNLFLSTQKQVBSYEGAVPLQDFGLANDTVR 1260

```

Db 1261 TKQHTAIFKKGREENLEGLGHQTKQIKVKYACTTRISNTSQNFVQTSKRALQFRL 1320
 Qy 743 ||||| 742
 Db 1321 PLEBTELEHRIIVDDTSTQMSQKHILTPSTLTQIDNREKGAITQSPJSDCIRESHSI 1380
 Qy 743 ||||| 742
 Db 1381 POANRSPLFAIAKVSFFSIRPIYITVLPQNSHSLPAASPKDKGCVQSSHFQGAKK 1440
 Qy 743 ||||| 742
 Db 1441 NUSLSALULEMTGQREWSGLGTSATNSVTKVQNTVLPKEDLPKSKVELLPKVIH 1500
 Qy 743 ||||| 742
 Db 1501 YOKULPPTESNGSPHLDVESSLGALTEGALKEANREPKYFPAIVATSSATPSK 1560
 Qy 743 ||||| 742
 Db 1561 LIAOPLANDHVGTOIPEEKMSQKRSPEKTAFAKQDITLSLANCENHAIANINGQWPF 1620
 Qy 743 ||||| 786
 Db 1621 EIRVPAKQGTBRLCSNPPFLPAHQKRIKTRTLAGQSHEDIDTQISYEMKEDPDII 1680
 Qy 787 DBENQSPSPQKTRHFAAVERLMDYGMSSSPHYLANAQSGSPQKVPKVPQBEFTD 846
 Db 1681 DBENQSPSPQKTRHFAAVERLMDYGMSSSPHYLANAQSGSPQKVPKVPQBEFTD 1740
 Qy 847 GSFTQPLVXGELNEHLLGGLFYIAEVEDNIMVTFMQAKSPFSYSSLSYEBDQGA 906
 Db 1741 GSFTQPLVXGELNEHLLGGLFYIAEVEDNIMVTFMQAKSPFSYSSLSYEBDQGA 1800
 Qy 907 EPRQHFYKNETKTYFMKVOHMAPTKDEFCAMATSDVLEKQVHSLGLGALVCHT 966
 Db 1801 EPRQHFYKNETKTYFMKVOHMAPTKDEFCAMATSDVLEKQVHSLGLGALVCHT 1860
 Qy 967 VYLANPAGROVTVQREAFPTTIDETKSNVPTENNENRCAPIOWHETPKENRFA 1026
 Db 1861 VYLANPAGROVTVQREAFPTTIDETKSNVPTENNENRCAPIOWHETPKENRFA 1920
 Qy 1027 INXYIMDTPLGLMAADORBWLVLSCSNENHSHFSGHVFVCKEYKVALNLYP 1086
 Db 1921 INXYIMDTPLGLMAADORBWLVLSCSNENHSHFSGHVFVCKEYKVALNLYP 1980
 Qy 1087 GUYFVWMLPSAGIARVLEHLAGHAGHTLFYFNSKQCPQAMASQNTDQBPOTAS 1146
 Db 1981 GUYFVWMLPSAGIARVLEHLAGHAGHTLFYFNSKQCPQAMASQNTDQBPOTAS 2040
 Qy 1147 GQYQNAKPLAKLHFKSTNINTEKQNTQYDILAPMILHSHKQACQNSHLSIQ 1206
 Db 2041 GQYQNAKPLAKLHFKSTNINTEKQNTQYDILAPMILHSHKQACQNSHLSIQ 2100
 Qy 1207 FTTHSGIATKQWCTYKGNSTGTGLWFGNDSGSIKNNFNPPILATITLHPTNITS 1266
 Db 2101 FTTHSGIATKQWCTYKGNSTGTGLWFGNDSGSIKNNFNPPILATITLHPTNITS 2160
 Qy 1267 TLMWLMGCLNCSMPLEGMESALSDAQITASITFPMFPAPEKSKYALQGSNMR 1326
 Db 2161 TLMWLMGCLNCSMPLEGMESALSDAQITASITFPMFPAPEKSKYALQGSNMR 2220
 Qy 1327 PQVNPWKEVDFQKTYKTOYQKSLLSYMYKFEPLSSQDQHWLPFGQKV 1386
 Db 2221 PQVNPWKEVDFQKTYKTOYQKSLLSYMYKFEPLSSQDQHWLPFGQKV 2280
 Qy 1387 QYQVNSGSPFVWNSLDPFLTYRLAHQSPQHQHQAIDMSEVGECAQY 1438
 Db 2281 QYQVNSGSPFVWNSLDPFLTYRLAHQSPQHQHQAIDMSEVGECAQY 2332

US-09-315-179-2
 / Sequence 2, Application US/09151179
 / Patent No. 6376463
 / GENERAL INFORMATION: John S
 / TITLE OF INVENTION: Modified Factor VIII
 / FILE REFERENCE: 75-9SH
 / CURRENT APPLICATION NUMBER: US/09/315-179
 / EARLIER APPLICATION NUMBER: U.S. 09/037,601
 / EARLIER FILING DATE: 1998-03-10
 / EARLIER APPLICATION NUMBER: U.S. 08/670,707
 / EARLIER FILING DATE: 1997-06-26
 / EARLIER APPLICATION NUMBER: PCT/US97/11155
 / EARLIER FILING DATE: 1997-06-26
 / EARLIER APPLICATION NUMBER: PCT/US94/13200
 / EARLIER FILING DATE: 1994-11-15
 / EARLIER APPLICATION NUMBER: U.S. 08/212,133
 / EARLIER FILING DATE: 1994-03-11
 / EARLIER APPLICATION NUMBER: U.S. 07/864,004
 / EARLIER FILING DATE: 1992-04-07
 / SOFTWARE: SEQ ID NO: 40
 / SOFTWARE: Patent In Ver. 2.0
 / SEQ ID NO 2
 / LENGTH: 2332
 / PRELIMINARY
 / COPIES: Homo sapiens

Query Match Similarity 94.1%; Score 7234; DB 4; Length 2332;
 Matches 1438; Conservative 0; Mismatches 0; Indels 894; Gaps 1;

Qy 1 ATATYTGAVELSDWDMQSDGLPVDNRPFPFPKSPFNTSVVTKTLFVFTVLEN 60
 Db 1 ATATYTGAVELSDWDMQSDGLPVDNRPFPFPKSPFNTSVVTKTLFVFTVLEN 60
 Qy 61 TAKPFPFMGLGFTTQAEYDVTITLQNAASHPVSLHAGVSYWAKSGAEYDDTQ 120
 Db 61 TAKPFPFMGLGFTTQAEYDVTITLQNAASHPVSLHAGVSYWAKSGAEYDDTQ 120
 Qy 121 REKEDKVPFGSGHTYVQWLKFNPMASDPLCTVTSYSLVDVLDVNSGLIGALVCR 180
 Db 121 REKEDKVPFGSGHTYVQWLKFNPMASDPLCTVTSYSLVDVLDVNSGLIGALVCR 180
 Qy 181 RGLSAKXTQTLHFTLLFVAPDQKQSHISTNSLMDQDAASAKAWPHVTVGNR 240
 Db 181 RGLSAKXTQTLHFTLLFVAPDQKQSHISTNSLMDQDAASAKAWPHVTVGNR 240
 Qy 241 SLPGLGCHKNSYVTVVGTPEVHSITFLGHTFLVNRHQSLSPETPTAQTLL 300
 Db 241 SLPGLGCHKNSYVTVVGTPEVHSITFLGHTFLVNRHQSLSPETPTAQTLL 300
 Qy 301 MDLQGLPCHISCHQDQKNAVYKUSCEPQPMQKNEEADYDDLDKSEMDVTFP 360
 Db 301 MDLQGLPCHISCHQDQKNAVYKUSCEPQPMQKNEEADYDDLDKSEMDVTFP 360
 Qy 361 DDNSPSFTQKCHQDQKNAVYKUSCEPQPMQKNEEADYDDLDKSEMDVTFP 420
 Db 361 DDNSPSFTQKCHQDQKNAVYKUSCEPQPMQKNEEADYDDLDKSEMDVTFP 420
 Qy 421 RKVKQKPMVWTFVNRHREKAKHSGEQTGLFVGVGHTLLIFPDQASPNLYPHOI 480
 Db 421 RKVKQKPMVWTFVNRHREKAKHSGEQTGLFVGVGHTLLIFPDQASPNLYPHOI 480
 Qy 481 TVDRPLSEBLKAKHKKOPPLVGLHITKATVTVDDSPDKLTRYTSFYVNE 540
 Db 481 TVDRPLSEBLKAKHKKOPPLVGLHITKATVTVDDSPDKLTRYTSFYVNE 540
 Qy 541 RDLASGLGPLLICITVQDQNGQMSDNRTLVFSFDEHRTNITENTQFPLFPAG 600
 Db 541 RDLASGLGPLLICITVQDQNGQMSDNRTLVFSFDEHRTNITENTQFPLFPAG 600

Qy 601 VOLBDFPQANMISINQVPSVPSLQVSLVLEHVAWYLLSIAQDQTLVIVSGTTPQI 660
 Db 601 VOLBDFPQANMISINQVPSVPSLQVSLVLEHVAWYLLSIAQDQTLVIVSGTTPQI 660
 Qy 661 MYVIEDTLVLPFPGSETVPMSEWQVPILOGNSDFPMRGMTALVKSICNDQYDTE 720
 Db 661 MYVIEDTLVLPFPGSETVPMSEWQVPILOGNSDFPMRGMTALVKSICNDQYDTE 720
 Qy 721 DSYEDISALSNKXNAIPERSF----- 742
 Db 721 DSYEDISALSNKXNAIPERSF----- 742
 Qy 743 ----- 760
 Db 743 ----- 760
 Qy 761 KXQVSSSLLMLLQSGFPFSGLSLSDQKACVETFDGSPSFOALUSNSINLSEHTFQ 840
 Db 761 KXQVSSSLLMLLQSGFPFSGLSLSDQKACVETFDGSPSFOALUSNSINLSEHTFQ 840
 Qy 743 ----- 742
 Db 743 ----- 742
 Qy 841 LHSQDNVFPFSGLOLALMEFLQTLVATLKLQLOFKVSGSTSNLSTIPDNLAJGTON 900
 Db 841 LHSQDNVFPFSGLOLALMEFLQTLVATLKLQLOFKVSGSTSNLSTIPDNLAJGTON 900
 Qy 743 ----- 742
 Db 743 ----- 742
 Qy 901 TBSLQSPVPMVPSQULTLVFKGKSPTVPSGSPLEBENNSKLLBESGLAMGSSBW 960
 Db 901 TBSLQSPVPMVPSQULTLVFKGKSPTVPSGSPLEBENNSKLLBESGLAMGSSBW 960
 Qy 743 ----- 742
 Db 743 ----- 742
 Qy 961 QNVASTEGEGLFKGRMHGIPALVYUNALFVSIILAKTKTNSATNRKTHIDGSL 1020
 Db 961 QNVASTEGEGLFKGRMHGIPALVYUNALFVSIILAKTKTNSATNRKTHIDGSL 1020
 Qy 743 ----- 742
 Db 743 ----- 742
 Qy 1021 LIENSFQWQKLESOTEFKVTPLDHRMLDMKQNTALRLHNSKNTSSKMNQOQ 1080
 Db 1021 LIENSFQWQKLESOTEFKVTPLDHRMLDMKQNTALRLHNSKNTSSKMNQOQ 1080
 Qy 743 ----- 742
 Db 743 ----- 742
 Qy 1081 KEGPTPDQAPDMSFFKOLFPSSARWQRTGEGSLNSQGSFQKQVSLGPEKSVG 1140
 Db 1081 KEGPTPDQAPDMSFFKOLFPSSARWQRTGEGSLNSQGSFQKQVSLGPEKSVG 1140
 Qy 743 ----- 742
 Db 743 ----- 742
 Qy 1141 QNFLGEGHGVYVGGSEFTQVGLKXENVPFSGRLFLTNLDLHNHTNQKCLQEBLK 1200
 Db 1141 QNFLGEGHGVYVGGSEFTQVGLKXENVPFSGRLFLTNLDLHNHTNQKCLQEBLK 1200
 Qy 743 ----- 742
 Db 743 ----- 742
 Qy 1201 KETLQBNVLVPGHVTGTRFKMLFLSTQRHVGSEYEGNAVLPQDFSIJNSTNR 1260
 Db 1201 KETLQBNVLVPGHVTGTRFKMLFLSTQRHVGSEYEGNAVLPQDFSIJNSTNR 1260
 Qy 743 ----- 742
 Db 743 ----- 742
 Qy 1261 TKGHTAHFKKGEENLGLGHQTKQIVKKTACTRISPTNSQNFVQTSKRALQFRL 1320
 Db 1261 TKGHTAHFKKGEENLGLGHQTKQIVKKTACTRISPTNSQNFVQTSKRALQFRL 1320
 Qy 743 ----- 742
 Db 743 ----- 742
 Qy 1321 PLESTELEGRIVDDTSQMSKNHLPSTLTQIDTNEKGAITQSPILSDCIRESHSI 1380
 Db 1321 PLESTELEGRIVDDTSQMSKNHLPSTLTQIDTNEKGAITQSPILSDCIRESHSI 1380
 Qy 743 ----- 742
 Db 743 ----- 742
 Qy 1381 PQANRSPJTAKVSPSPRIPVITVLPQUNSHLPAASTKDSQVSSSHFLQAKNK 1440
 Db 1381 PQANRSPJTAKVSPSPRIPVITVLPQUNSHLPAASTKDSQVSSSHFLQAKNK 1440
 Qy 743 ----- 742
 Db 743 ----- 742
 Qy 1441 NUSLSALVLEMTQDQREWSLQTSATNSVTKVYENTLPKEDLPKTSCKVLLPKVHI 1500
 Db 1441 NUSLSALVLEMTQDQREWSLQTSATNSVTKVYENTLPKEDLPKTSCKVLLPKVHI 1500
 Qy 743 ----- 742
 Db 743 ----- 742
 Qy 1501 YQKDLPTTENGSPGHLQVSGSLQGTGALWNEKPKYPLAVATSSAKTPSK 1560
 Db 1501 YQKDLPTTENGSPGHLQVSGSLQGTGALWNEKPKYPLAVATSSAKTPSK 1560
 Qy 743 ----- 742
 Db 743 ----- 742
 Qy 1561 LLDPLADHWHGTQVPEKNGQKSPKNTAKPKQUTLISJANCSNHAJANMEQKRF 1620
 Db 1561 LLDPLADHWHGTQVPEKNGQKSPKNTAKPKQUTLISJANCSNHAJANMEQKRF 1620
 Qy 743 ----- 786
 Db 743 ----- 786
 Qy 1621 ELVYVWAKQGTBLCONFVPLQKQKREKTRITLQSDQREIDYDITISVENKKEPDYI 1680
 Db 1621 ELVYVWAKQGTBLCONFVPLQKQKREKTRITLQSDQREIDYDITISVENKKEPDYI 1680
 Qy 787 DSDENQSPSPQKTRHFTFAVERLMDYGNSSSPHYLJNNAQSGSVQPKVVPQFTD 846
 Db 787 DSDENQSPSPQKTRHFTFAVERLMDYGNSSSPHYLJNNAQSGSVQPKVVPQFTD 846

Db 1681 DSDENQSPSPQKTRHFTFAVERLMDYGNSSSPHYLJNNAQSGSVQPKVVPQFTD 1740
 Qy 847 GSTFQPLVYRGELNHLGLGYPVTRAEVONTWTFPNQASRPYSFYSSLSYBEOQQA 906
 Db 847 GSTFQPLVYRGELNHLGLGYPVTRAEVONTWTFPNQASRPYSFYSSLSYBEOQQA 906
 Qy 1741 GSTFQPLVYRGELNHLGLGYPVTRAEVONTWTFPNQASRPYSFYSSLSYBEOQQA 1800
 Db 1741 GSTFQPLVYRGELNHLGLGYPVTRAEVONTWTFPNQASRPYSFYSSLSYBEOQQA 1800
 Qy 907 EPKQFVYKNEKTKTFKVOGHNAVPOKDFCKANAYTSVDOLKOVHSGIIGPLLCHT 966
 Db 907 EPKQFVYKNEKTKTFKVOGHNAVPOKDFCKANAYTSVDOLKOVHSGIIGPLLCHT 966
 Qy 1801 EPKQFVYKNEKTKTFKVOGHNAVPOKDFCKANAYTSVDOLKOVHSGIIGPLLCHT 1860
 Db 1801 EPKQFVYKNEKTKTFKVOGHNAVPOKDFCKANAYTSVDOLKOVHSGIIGPLLCHT 1860
 Qy 967 WTNAPAGROVVOGEPALFFITFDETKSYMTFENNERKCARCQIMQDFPKENYFHA 1026
 Db 967 WTNAPAGROVVOGEPALFFITFDETKSYMTFENNERKCARCQIMQDFPKENYFHA 1026
 Qy 1861 WTNAPAGROVVOGEPALFFITFDETKSYMTFENNERKCARCQIMQDFPKENYFHA 1920
 Db 1861 WTNAPAGROVVOGEPALFFITFDETKSYMTFENNERKCARCQIMQDFPKENYFHA 1920
 Qy 1027 INXTIMDTLQVLAQDQORTWYLLSLNSGNSNTSHSIFSGRFTVKKKERYKMAVLYP 1086
 Db 1027 INXTIMDTLQVLAQDQORTWYLLSLNSGNSNTSHSIFSGRFTVKKKERYKMAVLYP 1086
 Qy 1921 INXTIMDTLQVLAQDQORTWYLLSLNSGNSNTSHSIFSGRFTVKKKERYKMAVLYP 1980
 Db 1921 INXTIMDTLQVLAQDQORTWYLLSLNSGNSNTSHSIFSGRFTVKKKERYKMAVLYP 1980
 Qy 1087 OVYFVEMVLSKADTHRVFCLIGBLHAGMCLFLVYNSKQCTPLQANSHIRPQITAS 1146
 Db 1087 OVYFVEMVLSKADTHRVFCLIGBLHAGMCLFLVYNSKQCTPLQANSHIRPQITAS 1146
 Qy 1981 OVYFVEMVLSKADTHRVFCLIGBLHAGMCLFLVYNSKQCTPLQANSHIRPQITAS 2040
 Db 1981 OVYFVEMVLSKADTHRVFCLIGBLHAGMCLFLVYNSKQCTPLQANSHIRPQITAS 2040
 Qy 1147 QYQVAPKALBKLKXGSEINAKSTKSPWVTLQVLAQMIHGTQKQAKQFSSLYISQ 1206
 Db 1147 QYQVAPKALBKLKXGSEINAKSTKSPWVTLQVLAQMIHGTQKQAKQFSSLYISQ 1206
 Qy 2041 QYQVAPKALBKLKXGSEINAKSTKSPWVTLQVLAQMIHGTQKQAKQFSSLYISQ 2100
 Db 2041 QYQVAPKALBKLKXGSEINAKSTKSPWVTLQVLAQMIHGTQKQAKQFSSLYISQ 2100
 Qy 1207 FTKMGEKQK 1266
 Db 1207 FTKMGEKQK 1266
 Qy 2101 FTKMGEKQK 2160
 Db 2101 FTKMGEKQK 2160
 Qy 1267 FTKMGEKQK 1322
 Db 1267 FTKMGEKQK 1322
 Qy 1327 FTKMGEKQK 1386
 Db 1327 FTKMGEKQK 1386
 Qy 2221 FTKMGEKQK 2280
 Db 2221 FTKMGEKQK 2280
 Qy 1387 KYQVQK 1438
 Db 1387 KYQVQK 1438
 Qy 2381 KYQVQK 2332
 Db 2381 KYQVQK 2332

RESUME 13
 US-09-523-656-2
 ; Sequence 2, Application US/09523656
 ; Patent No. 6458563
 ; GENERAL INFORMATION:
 ; INVENTOR: John
 ; TITLE OF INVENTION: MODIFIED FACTOR VIII
 ; FILE REFERENCE: 75-951
 ; CURRENT APPLICATION NUMBER: US/09/523,656
 ; EARLIER APPLICATION NUMBER: US/09/037,601
 ; EARLIER FILING DATE: 1998-03-10
 ; EARLIER APPLICATION NUMBER: 08/670,707
 ; EARLIER FILING DATE: 1996-06-26
 ; NAME OF APPLICANT: Genzyme Corporation
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 2332
 ; ORGANISM: Homo sapiens
 US-09-523-656-2

Query Match 94.1%; Score 7214; DB 4; Length 2332;
 Similarity 94.1%;
 Matches 1438; Conservative 0; Mismatches 0; Indels 894; Gaps 1;
 Qy 1 ATNRYTGLAVELSHQSDSGELPVDQAPFPVPSKPPFNSTVTKKTLVPEVTLN 60
 Db 1 ATNRYTGLAVELSHQSDSGELPVDQAPFPVPSKPPFNSTVTKKTLVPEVTLN 60

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2003, 17:03:15 ; Search time 43 Seconds
(without alignments)
6219.640 Million cell updates/sec

- ```

Perfect score: 7631
Sequence: 1 ATTTTATGAVLHSDWQSD.....WHQIALEHVEVCQAQDLY 1438

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA*
1: /cgn2/6/prodata1/pubpaa/US07_FUECOMB.pep.*
2: /cgn2/6/prodata1/pubpaa/ECT_NEW_PUB.pep.*
3: /cgn2/6/prodata1/pubpaa/US06_FUECOMB.pep.*
4: /cgn2/6/prodata1/pubpaa/US06_FUECOMB.pep.*
5: /cgn2/6/prodata1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2/6/prodata1/pubpaa/ECTUS_FUECOMB.pep.*
7: /cgn2/6/prodata1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2/6/prodata1/pubpaa/US08_NEW_PUB.pep.*
9: /cgn2/6/prodata1/pubpaa/US09_FUECOMB.pep.*
10: /cgn2/6/prodata1/pubpaa/US09_FUECOMB.pep.*
11: /cgn2/6/prodata1/pubpaa/US09C_FUECOMB.pep.*
12: /cgn2/6/prodata1/pubpaa/US09C_FUECOMB.pep.*
13: /cgn2/6/prodata1/pubpaa/US10_FUECOMB.pep.*
14: /cgn2/6/prodata1/pubpaa/US10B_FUECOMB.pep.*
15: /cgn2/6/prodata1/pubpaa/US10B_FUECOMB.pep.*
16: /cgn2/6/prodata1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2/6/prodata1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2/6/prodata1/pubpaa/US00_FUECOMB.pep.*
19: /cgn2/6/prodata1/pubpaa/US00_FUECOMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID                | Description             |
|------------|-------|-------------|--------|-------------------|-------------------------|
| 1          | 7891  | 100         | 438    | US-10-047-297-1   | Sequence 1, April 1978  |
| 2          | 7891  | 100         | 438    | US-10-047-297-1   | Sequence 1, April 1978  |
| 3          | 7691  | 100         | 438    | US-10-255-900-1   | Sequence 2, April 1978  |
| 4          | 7674  | 99.8        | 441    | US-10-095-718-2   | Sequence 2, April 1978  |
| 5          | 7674  | 99.8        | 441    | US-10-095-718-2   | Sequence 2, April 1978  |
| 6          | 7234  | 94.1        | 2332   | US-10-131-5103-2  | Sequence 2, April 1978  |
| 7          | 7234  | 94.1        | 2332   | US-10-131-5103-2  | Sequence 2, April 1978  |
| 8          | 7234  | 94.1        | 2332   | US-10-187-319-2   | Sequence 2, April 1978  |
| 9          | 7227  | 94.0        | 2331   | US-10-133-907-4   | Sequence 4, April 1978  |
| 10         | 7227  | 94.0        | 2331   | US-10-133-907-4   | Sequence 4, April 1978  |
| 11         | 6599  | 85.8        | 4431   | US-10-095-718-2   | Sequence 4, April 1978  |
| 12         | 6489  | 84.8        | 4432   | US-10-131-5108-39 | Sequence 39, April 1978 |
| 13         | 6489  | 84.8        | 4432   | US-10-131-5108-39 | Sequence 39, April 1978 |
| 14         | 6227  | 81.0        | 2133   | US-10-131-5103-37 | Sequence 37, April 1978 |
| 15         | 6227  | 81.0        | 2133   | US-10-187-319-37  | Sequence 37, April 1978 |

## ALIGNMENTS

## RESULT 1

```

US-10-006-091-1 1 APPLICATION US/1006091
Pub No US 10200910273001
GENERAL INFORMATION:
APPLICANT: Cho, Myung-Sam
APPLICANT: Chan, Sham-Yuen
APPLICANT: Kaisey, William
APPLICANT: Kasey, William
FILE OF INVENTION: Expression System for
TITLE REFERENCE: MSB-7255.1
CURRENT APPLICATION NUMBER: US/1006,091
CURRENT FILING DATE: 2001-12-06
NUMBER OF SEQ ID NOS: 2
SEQ ID NO 1 Parentin Ver. 2.0
LENGTH: 1438
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
GENERAL INFORMATION: Description of Artifici
GENERAL INFORMATION: human factor VIII segm
US-10-006-091-1

```

|    | Query Match                                                  | 100.0%       | Score 7691; | DB 14;     | Length 1438; |
|----|--------------------------------------------------------------|--------------|-------------|------------|--------------|
|    | Best Local Similarity                                        | 100.0%;      | Pred. No.   | Mismatches | Gaps         |
|    | Matches 1438;                                                | Conservative | 0;          | Indels     | 0;           |
| OY | 1 ATATTTGAGLASHMDSGDLGVADPPKPPSPPTSVVTKLVETVFHLEIN 60        |              |             |            |              |
| DB | 1 ATATTGAGLASHMDSGDLGVADPPKPPSPPTSVVTKLVETVFHLEIN 60         |              |             |            |              |
| OY | 61 TAAKPFPWMLGLTQAEVTFTVTITLNKAHPSHVAHVSTWAASAKFDIQSQ 120    |              |             |            |              |
| DB | 61 TAAKPFPWMLGLTQAEVTFTVTITLNKAHPSHVAHVSTWAASAKFDIQSQ 120    |              |             |            |              |
| OY | 121 AKRHHDTVPFGSHTTIVQLNKGPSAPSDPLTSTLTSLANDVMSGLALGATVC 180 |              |             |            |              |
| DB | 121 AKRHHDTVPFGSHTTIVQLNKGPSAPSDPLTSTLTSLANDVMSGLALGATVC 180 |              |             |            |              |





```

Db 721 DEYDIDYANVLKGRMALEPRFSRQNPVLVQRKQREITRTTLOQSGREIDYDIDISVENKX 780
Qy 781 EDYDIDYDENOQSPSFOKTRHFTFAAVERLMDYGNSSPHVLANKAQSGVQPKVY 840
Db 781 EDYDIDYDENOQSPSFOKTRHFTFAAVERLMDYGNSSPHVLANKAQSGVQPKVY 840
Qy 841 FOFYDGSFTQPLRGELNHEHLLGLGPIBAEDVNIWTFRQASRPYSYSSLSVEE 900
Db 841 FOFYDGSFTQPLRGELNHEHLLGLGPIBAEDVNIWTFRQASRPYSYSSLSVEE 900
Qy 901 DORQAGSRPHQVKNETNTYFKVKORHMAPTKDEPOCKANAFSDVLKQWBSGLGP 960
Db 901 DORQAGSRPHQVKNETNTYFKVKORHMAPTKDEPOCKANAFSDVLKQWBSGLGP 960
Qy 961 LVAUHTNINPANGQVTVQEPALFTTFFDTSKSTFTFEMENENCAPON QWEDPTFE 1020
Db 961 LVAUHTNINPANGQVTVQEPALFTTFFDTSKSTFTFEMENENCAPON QWEDPTFE 1020
Qy 1021 NYREHAIKNTYMOITPOLMAQDOORIKWLLAKNGSNRHHIFHFGVTFVAKKRYDIA 1080
Db 1021 NYREHAIKNTYMOITPOLMAQDOORIKWLLAKNGSNRHHIFHFGVTFVAKKRYDIA 1080
Qy 1081 LKNVQGVTFVTRKAGIWRVCELCXHUHAKQSTPLVYNSQCPKQKQASHGRTD 1140
Db 1081 LKNVQGVTFVTRKAGIWRVCELCXHUHAKQSTPLVYNSQCPKQKQASHGRTD 1140
Qy 1141 FOYTASQGVQKQALVHLYKQVGNSTQSGTQKQVLLAKNTHUWQKQKQKQ 1200
Db 1141 FOYTASQGVQKQALVHLYKQVGNSTQSGTQKQVLLAKNTHUWQKQKQKQ 1200
Qy 1201 SHYQOFTINVLQKQKQVGNSTQGVWYFGVQKQSGKNNIFRPLIANYIRLEPT 1260
Db 1201 SHYQOFTINVLQKQKQVGNSTQGVWYFGVQKQSGKNNIFRPLIANYIRLEPT 1260
Qy 1261 HYSITSTLMEJLQCLNCSMPLOWESKALQAOITNASTFTFNPATWSPSIRLHLQ 1320
Db 1261 HYSITSTLMEJLQCLNCSMPLOWESKALQAOITNASTFTFNPATWSPSIRLHLQ 1320
Qy 1321 RNSARFQONWPEKQVLDFOKMYKQVTVTQVQKSLTSMYKSTFISBQOQKQWTLF 1380
Db 1321 RNSARFQONWPEKQVLDFOKMYKQVTVTQVQKSLTSMYKSTFISBQOQKQWTLF 1380
Qy 1381 FONGKATVQKQDSFTFVNSISDPLATYLARHPQSNVQKQALMEVCEQOQL 1438
Db 1381 FONGKATVQKQDSFTFVNSISDPLATYLARHPQSNVQKQALMEVCEQOQL 1438

RESULT 4
US-10-047-257-1
; Sequence 2, Application US/10095718
; Publication No. US2002013156A1
; GENERAL INFORMATION:
; APPLICANT: Chao, Hengjun
; APPLICANT: Chao, Hengjun
; APPLICANT: Burslein, Haim
; APPLICANT: Lynch, Carmel
; APPLICANT: Kitepen, Kaiti
; APPLICANT: Kitepen, Kaiti
; TITLE OF INVENTOR: Adeno-Associated Virus Vectors Encoding Factor VIII and
; FILE REFERENCE: 3502/204375 US/10/095,718
; CURRENT FILING DATE: 2002-03-12
; PRIOR FILING DATE: 2001-06-22
; PRIOR FILING DATE: 2001-06-22
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; SEQ ID NO 2
; TYPE: PAT
; ORGANISM: Homo sapiens B-domain deleted factor VIII

```

```

; FEATURE:
; INFORMATION: Homo sapiens BDD FVIII
US-10-095-718-2

```

```

Query Match 99.88; Score 7674; DB 14; Length 1474;
Seq. Identity 99.88; Pct. Id. 99.88;
Match 1188; Conservative 0; Indels 14; Gaps 1;

```

```

Qy 1 ATATYTLGAVLSMDMDAGDGLSELPVDARFPVPPVPSFPFTSVYKTKLVEFTVHLN 60
Db 20 ATATYTLGAVLSMDMDAGDGLSELPVDARFPVPPVPSFPFTSVYKTKLVEFTVHLN 79
Qy 61 ATKRTFPMWGLGFTQAEYVTVVTLKQNAASPEVLSHAGVSPKASGCAEVDQDSQ 120
Db 80 INKAPFMWGLGFTQAEYVTVVTLKQNAASPEVLSHAGVSPKASGCAEVDQDSQ 139
Qy 121 REKEDDVTFFGSGSHYTMVLKENGPMASDPLCLTYSLSHVDLVKQNSGLIALLVC 180
Db 140 REKEDDVTFFGSGSHYTMVLKENGPMASDPLCLTYSLSHVDLVKQNSGLIALLVC 199
Qy 181 KEGAAKQVTLRFTLFLVAPDQSGHSPSTQSLMQRDQASANRAMPQHTVQYVNR 240
Db 200 KEGAAKQVTLRFTLFLVAPDQSGHSPSTQSLMQRDQASANRAMPQHTVQYVNR 259
Qy 241 SLPLTLCGRKSYVYHVGMGTVPVHVSFLGHTFTFLVNRQASLSISPTFTATQALL 300
Db 260 SLPLTLCGRKSYVYHVGMGTVPVHVSFLGHTFTFLVNRQASLSISPTFTATQALL 319
Qy 301 MDGLQPLPCHLSHHQDQHEAVYKVSCEPSPQKMNQWNEEADYDDLTQSDNDVRF 360
Db 320 MDGLQPLPCHLSHHQDQHEAVYKVSCEPSPQKMNQWNEEADYDDLTQSDNDVRF 379
Qy 361 DDNDSQSFQIASHAKKHPKWHYTAASEEDQVAPLADPDRSKSYQVLANGQK 420
Db 380 DDNDSQSFQIASHAKKHPKWHYTAASEEDQVAPLADPDRSKSYQVLANGQK 439
Qy 421 KKKYKVPKNTDFTKTRALQKSGGLGLPQKSGVQTLIIFKQNAQSRPHTVPHQ 480
Db 440 KKKYKVPKNTDFTKTRALQKSGGLGLPQKSGVQTLIIFKQNAQSRPHTVPHQ 499
Qy 481 TQVRSYKRSKQVHAKDPTLQSEKPKYKNTVYKQVYKQSDPCLKRYKVS NVMS 540
Db 500 TQVRSYKRSKQVHAKDPTLQSEKPKYKNTVYKQVYKQSDPCLKRYKVS NVMS 559
Qy 541 RLKASGLLQPLLCYKESVDQKQNMQRKRVILFVSFBNRSNYLTIQRTLPNAG 600
Db 560 RLKASGLLQPLLCYKESVDQKQNMQRKRVILFVSFBNRSNYLTIQRTLPNAG 619
Qy 601 VQLEDQFPKASNIWHISQIVYNSQLVCLHVAWVILISGAQDPLSFSGFTFH 660
Db 620 VQLEDQFPKASNIWHISQIVYNSQLVCLHVAWVILISGAQDPLSFSGFTFH 679
Qy 661 NVYVTDLTLPFPGSGTFVMSNPLMGLQCNDFPKNGMALKLVSCDQDGYTE 720
Db 680 NVYVTDLTLPFPGSGTFVMSNPLMGLQCNDFPKNGMALKLVSCDQDGYTE 739
Qy 721 DSDYDANLKNNAIFSPSSPNQ-----PVLKQREITRTTLOQSGREIT 766
Db 740 DSDYDANLKNNAIFSPSSPNQSHRPSKQKQVTFVVLKQREITRTTLOQSGRE 799
Qy 767 ETDYDIDYVENKEDFOITDENOQSPSFOKTRHFTFAAVERLMDYGNSSPHVLN 826
Db 800 ETDYDIDYVENKEDFOITDENOQSPSFOKTRHFTFAAVERLMDYGNSSPHVLN 859
Qy 827 BAQSGVQPKVQKQVTFQDSFTQPLRGELNHEHLLGLGPIBAEDVNIWTFRQAS 886
Db 860 BAQSGVQPKVQKQVTFQDSFTQPLRGELNHEHLLGLGPIBAEDVNIWTFRQAS 919
Qy 887 RPYTSYSSLSVEEDQKQKPRKPKYKQNTYTKYKVKORHMAPTKDEPOCKANAFSD 946
Db 920 RPYTSYSSLSVEEDQKQKPRKPKYKQNTYTKYKVKORHMAPTKDEPOCKANAFSD 979
Qy 947 VQLEDVHSGSLGLPVLCHTNINPANGQVTVQEPALFTTFFDTSKSTFTFEMEN 1006

```



[illegible]

```

RESULT 5
US-10-131-510A-2
1 SEQUENCE 2, Application US/10131510A
2 INVENTOR: KOLLAR, JOHN S.
3 APPLICATION: US20030146555A1
4
5 TITLE OF INVENTION: Modified Factor VIII
6
7 CURRENT REFERENCE: 75-350
8 CURRENT FILING DATE: 2002-07-10
9
10 PRIOR APPLICATION NUMBER: U.S. 09/315,179
11 PRIOR FILING DATE: 1999-05-20
12
13 PRIOR APPLICATION NUMBER: U.S. 09/037,601
14 PRIOR FILING DATE: 1998-03-10
15
16 PRIOR APPLICATION NUMBER: U.S. 08/670,707
17 PRIOR FILING DATE: 1996-06-26
18
19 PRIOR FILING DATE: 1996-06-26
20 PRIOR FILING DATE: 1996-06-26/0897/11185
21
22 PRIOR APPLICATION NUMBER: US94/13200
23 PRIOR FILING DATE: 1994-11-15
24
25 PRIOR APPLICATION NUMBER: U.S. 08/212,133
26 PRIOR FILING DATE: 1994-01-25
27
28 PRIOR APPLICATION NUMBER: U.S. 07/864,004
29 PRIOR FILING DATE: 1992-04-07
30
31 NUMBER OF SEQ ID NOS: 40
32
33 SEQ ID NO 2: Parentin Ver. 2.0
34
35 LENGTH: 2332
36
37 TYPE: PRT
38
39 US-10-131-510A-2
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949

```

|                            |        |               |                                                               |              |
|----------------------------|--------|---------------|---------------------------------------------------------------|--------------|
| Query Match                | 94.1%; | Score 7234;   | DB 12;                                                        | Length 2332; |
| Best Local Similarity      | 61.7%; | Pfam No. 0;   |                                                               |              |
| Matches 1438; Conservative |        | 0; Mismatches | 0; Indels                                                     | 894; Gaps    |
| QY                         |        | 1             | ATRTYTLGAGVLSWDSWDQSDGLPELDYDARPPRPVKSFPPFTSVVYKTKLTFEETVALFN | 60           |

Query Match



Db 541 |||||GL||P||L||C||K||E||S||V||Q||R||Q||N||S||D||G||N||V||L||F||S||V||P||O||R||N||S||V||T||E||N||I||O||R||L||F||N||P||A||Q 600  
 Qy 601 V||O||L||E||D||P||O||A||N||T||N||G||V||P||D||S||Q||L||Q||V||A||J||E||V||A||V||T||L||E||C||A||D||E||L||F||A||V||P||E||C||T||O||K||I 660  
 Db 601 V||O||L||E||D||P||O||A||N||T||N||G||V||P||D||S||Q||L||Q||V||A||J||E||V||A||V||T||L||E||C||A||D||E||L||F||A||V||P||E||C||T||O||K||I 660  
 Qy 661 K||V||E||D||T||L||F||P||E||S||C||E||T||V||N||G||P||D||M||L||O||C||H||N||S||P||O||R||N||S||O||M||A||L||K||S||G||O||K||O||W||V||E|| 720  
 Db 661 K||V||E||D||T||L||F||P||E||S||C||E||T||V||N||G||P||D||M||L||O||C||H||N||S||P||O||R||N||S||O||M||A||L||K||S||G||O||K||O||W||V||E|| 720  
 Qy 721 P||E||D||I||G||A||V||L||S||O||N||N||A||P||E||R||P|| 742  
 Db 721 P||E||D||I||G||A||V||L||S||O||N||N||A||P||E||R||P|| 742  
 Qy 743 ----- 742  
 Db 761 K||I||Q||V||S||E||L||L||K||A||Q||P||H||G||L||S||E||L||Q||A||T||E||N||F||S||D||P||S||C||A||L||D||N||S||L||E||N||T||H||E||P||Q 840  
 Qy 761 K||I||Q||V||S||E||L||L||K||A||Q||P||H||G||L||S||E||L||Q||A||T||E||N||F||S||D||P||S||C||A||L||D||N||S||L||E||N||T||H||E||P||Q 840  
 Db 783 ----- 742  
 Qy 841 L||H||S||G||M||V||T||P||E||S||Q||L||E||N||K||E||L||G||T||A||N||T||E||L||K||D||F||K||V||S||T||S||N||L||I||S||T||P||D||N||L||A||G||T||N 900  
 Db 841 L||H||S||G||M||V||T||P||E||S||Q||L||E||N||K||E||L||G||T||A||N||T||E||L||K||D||F||K||V||S||T||S||N||L||I||S||T||P||D||N||L||A||G||T||N 900  
 Qy 743 ----- 742  
 Db 901 T||E||S||L||G||P||S||P||V||H||O||L||D||T||L||E||G||K||S||J||E||S||S||Q||P||L||S||E||R||N||S||K||L||E||S||G||L||O||S||Q||E||S||H 960  
 Qy 901 T||E||S||L||G||P||S||P||V||H||O||L||D||T||L||E||G||K||S||J||E||S||S||Q||P||L||S||E||R||N||S||K||L||E||S||G||L||O||S||Q||E||S||H 960  
 Db 943 ----- 742  
 Qy 943 G||O||N||V||S||E||S||G||L||E||K||E||A||N||G||P||A||L||T||O||M||A||L||K||V||S||L||A||K||N||K||S||N||S||A||T||R||K||H||D||E||S||L 1020  
 Db 943 ----- 742  
 Qy 1021 L||E||N||S||P||V||Q||L||E||S||T||E||P||K||Y||T||L||H||O||R||M||A||D||O||N||F||A||L||A||H||N||G||K||T||S||S||K||E||W||O||O||K 1080  
 Db 1021 L||E||N||S||P||V||Q||L||E||S||T||E||P||K||Y||T||L||H||O||R||M||A||D||O||N||F||A||L||A||H||N||G||K||T||S||S||K||E||W||O||O||K 1080  
 Qy 743 ----- 742  
 Db 1081 K||E||G||P||D||D||Q||M||S||P||F||K||M||F||L||P||S||R||A||N||T||O||R||H||G||N||S||L||Q||S||Q||S||P||Q||V||S||G||P||E||S||F||E||G 1140  
 Qy 743 ----- 742  
 Db 1141 Q||N||F||E||K||N||V||K||G||E||P||T||O||V||G||L||K||E||M||V||P||S||S||N||L||F||N||L||M||E||N||T||H||O||E||K||I||O||B||E||E||K 1200  
 Qy 743 ----- 742  
 Db 1201 K||E||T||L||O||R||N||V||L||P||O||H||T||V||G||T||K||N||F||K||M||L||A||S||T||Q||N||V||E||S||T||E||G||A||V||P||L||Q||F||S||L||N||D||S||T||R 1260  
 Qy 743 ----- 742  
 Db 1261 T||O||K||T||A||N||F||K||E||H||E||N||L||G||O||N||K||O||V||E||K||A||T||T||R||I||S||P||T||S||Q||N||V||P||O||R||S||K||A||L||Q||P||L 1320  
 Qy 743 ----- 742  
 Db 1321 P||A||B||E||T||E||K||R||I||V||D||O||T||S||O||N||K||M||L||P||S||T||Q||D||N||E||K||E||K||A||T||O||S||P||L||C||D||I||E||S||H||S||I 1380  
 Qy 743 ----- 742  
 Db 1381 P||O||A||N||S||P||L||A||K||S||F||E||S||I||R||P||Y||T||V||T||V||F||O||N||S||H||L||P||A||S||Y||K||S||G||O||V||S||H||L||Q||A||N||K 1440  
 Qy 743 ----- 742  
 Db 1441 N||U||S||L||A||L||E||M||T||G||O||R||S||L||G||T||S||N||V||T||K||Y||E||N||T||L||F||E||D||L||P||K||S||G||K||V||E||L||L||P||K||H|| 1500  
 Qy 743 ----- 742  
 Db 1501 Y||O||K||D||L||P||T||S||N||G||H||L||D||V||E||S||L||Q||T||E||G||A||K||N||E||N||R||K||G||V||F||A||V||T||E||S||S||A||K||T||E||S||K 1560  
 Qy 743 ----- 742  
 Db 1561 L||A||D||P||A||N||H||V||G||L||P||K||E||W||S||Q||R||S||P||E||K||A||P||K||O||T||L||S||I||N||C||H||A||A||I||N||E||Q||N||F 1620  
 Qy 743 -----S||N||P||V||A||H||O||R||I||N||T||L||Q||S||E||I||D||D||T||S||V||E||K||E||D||D||Y 1680

Db 1621 E||E||V||T||W||A||Q||G||E||T||E||L||C||S||Q||H||P||V||A||J||Q||E||R||T||T||L||Q||S||Q||E||I||D||Y||D||T||S||V||E||N||K||E||D||D||Y 1680  
 Qy 767 D||E||R||O||S||P||E||S||P||O||K||A||T||R||H||F||T||A||N||E||R||M||A||D||Y||O||M||S||S||H||P||V||L||E||N||A||S||G||S||V||P||K||V||V||E||R||D|| 846  
 Db 1661 D||E||R||O||S||P||E||S||P||O||K||A||T||R||H||F||T||A||N||E||R||M||A||D||Y||O||M||S||S||H||P||V||L||E||N||A||S||G||S||V||P||K||V||V||E||R||D|| 1740  
 Qy 847 G||S||T||O||P||Y||T||E||G||L||N||E||H||L||G||L||P||T||A||N||E||D||I||N||M||V||T||P||O||A||S||P||S||P||Y||S||S||L||S||E||D||O||R||Q||A 906  
 Db 1741 G||S||T||O||P||Y||T||E||G||L||N||E||H||L||G||L||P||T||A||N||E||D||I||N||M||V||T||P||O||A||S||P||S||P||Y||S||S||L||S||E||D||O||R||Q||A 1800  
 Qy 907 E||P||K||F||A||Z||O||N||E||T||F||K||V||Q||H||R||M||A||P||K||O||S||F||C||A||M||A||W||F||S||D||V||E||L||E||H||S||G||L||P||L||U||C||H|| 966  
 Db 1801 E||P||K||F||A||Z||O||N||E||T||F||K||V||Q||H||R||M||A||P||K||O||S||F||C||A||M||A||W||F||S||D||V||E||L||E||H||S||G||L||P||L||U||C||H|| 1860  
 Qy 967 N||T||A||N||P||A||G||O||V||T||Y||Q||E||A||F||F||T||E||D||K||S||W||V||T||E||N||E||N||S||C||A||P||C||N||Q||M||E||D||P||T||E||N||T||R||E||H||A 1026  
 Db 1861 N||T||A||N||P||A||G||O||V||T||Y||Q||E||A||F||F||T||E||D||K||S||W||V||T||E||N||E||N||S||C||A||P||C||N||Q||M||E||D||P||T||E||N||T||R||E||H||A 1920  
 Qy 1027 I||N||G||Y||M||O||T||L||P||Q||L||W||A||Q||O||R||I||W||L||L||S||M||G||S||N||E||N||S||I||H||F||S||G||V||T||V||K||E||B||Y||O||A||L||Y||N||P 1086  
 Db 1921 I||N||G||Y||M||O||T||L||P||Q||L||W||A||Q||O||R||I||W||L||L||S||M||G||S||N||E||N||S||I||H||F||S||G||V||T||V||K||E||B||Y||O||A||L||Y||N||P 1980  
 Qy 1087 O||V||E||T||E||M||P||S||A||G||I||V||V||E||L||L||G||B||H||A||H||Q||M||S||T||L||F||A||V||S||K||O||T||P||L||O||A||S||H||O||R||P||O||I||T||A||S 1146  
 Db 1981 O||V||E||T||E||M||P||S||A||G||I||V||V||E||L||L||G||B||H||A||H||Q||M||S||T||L||F||A||V||S||K||O||T||P||L||O||A||S||H||O||R||P||O||I||T||A||S 2040  
 Qy 1147 Q||O||G||O||N||A||P||K||A||L||H||V||E||L||L||G||B||H||A||H||Q||M||S||T||L||F||A||V||S||K||O||T||P||L||O||A||S||H||O||R||P||O||I||T||A||S 1206  
 Db 2041 Q||O||G||O||N||A||P||K||A||L||H||V||E||L||L||G||B||H||A||H||Q||M||S||T||L||F||A||V||S||K||O||T||P||L||O||A||S||H||O||R||P||O||I||T||A||S 2100  
 Qy 1207 F||I||N||Y||S||L||D||O||K||O||N||T||G||R||E||S||T||G||T||A||F||F||O||N||S||O||S||G||K||Y||N||E||I||A||N||V||I||A||R||V||I||A||P||H||V||S||I||S 1266  
 Db 2101 F||I||N||Y||S||L||D||O||K||O||N||T||G||R||E||S||T||G||T||A||F||F||O||N||S||O||S||G||K||Y||N||E||I||A||N||V||I||A||R||V||I||A||P||H||V||S||I||S 2160  
 Qy 1267 T||E||M||K||H||O||C||L||S||C||M||P||C||H||E||N||K||I||S||D||A||Q||T||S||T||F||F||N||V||T||W||S||S||A||R||H||L||Q||S||N||A||R 1326  
 Db 2161 T||E||M||K||H||O||C||L||S||C||M||P||C||H||E||N||K||I||S||D||A||Q||T||S||T||F||F||N||V||T||W||S||S||A||R||H||L||Q||S||N||A||R 2220  
 Qy 1327 P||O||A||N||P||C||H||L||O||V||D||O||N||K||Y||T||G||V||T||O||G||K||L||S||T||M||Y||K||E||L||S||S||O||D||O||H||W||L||F||P||O||N||K 1386  
 Db 2221 P||O||A||N||P||C||H||L||O||V||D||O||N||K||Y||T||G||V||T||O||G||K||L||S||T||M||Y||K||E||L||S||S||O||D||O||H||W||L||F||P||O||N||K 2280  
 Qy 1387 K||Y||T||O||G||S||P||P||V||N||S||L||P||L||I||A||R||I||K||H||I||P||O||S||V||H||Q||A||L||M||E||V||G||C||E||Q||A||Y 1438  
 Db 2281 K||Y||T||O||G||S||P||P||V||N||S||L||P||L||I||A||R||I||K||H||I||P||O||S||V||H||Q||A||L||M||E||V||G||C||E||Q||A||Y 2332

## RESULTS

US-10-187-319-2

; Sequence 2, Application US/10187319  
 ; Publication No. US2003006875A1  
 ; GENERAL INFORMATION:  
 ; INVENTOR: K. John S.  
 ; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII  
 ; NUMBER OF SEQUENCES: 40  
 ; CORRESPONDENCE ADDRESS:  
 ; Dr. John S. Kohn  
 ; STREET: 5370 Radcliff Circle Suite 201  
 ; CITY: Boulder  
 ; STATE: Colorado  
 ; COUNTRY: USA  
 ; ZIP: 80503  
 ; COMPUTER READABLE FORM:  
 ; US-10-187-319-2  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; SOFTWARE: SYSTEMS PC-DOS/MS-DOS  
 ; SOFTWARE: SYSTEMS PC-DOS/MS-DOS  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/10187319  
 ; FILING DATE: 27-Aug-2002  
 ; CROSS-REFERENCE TO OTHER APPLICATIONS:  
 ; PRIORITY APPLICATION DATA:  
 ; APPLICATION NUMBER: US 09/523,656





QY 743 ----- 742  
 Db 1160 QNFASRKKVVGGEFTDVLGEMVFPSSNFIPTNLDLHNTRQKCKIOETEK 1219  
 QY 743 ----- 742  
 Db 1220 KETLIDENVLQIHTYTGTINPHNLSTFSTRNGSFGAIPALQDFRSLNDSTNR 1279  
 QY 743 ----- 742  
 Db 1280 TCKTAFHSKXGEENLEGANQTKQVKEVACTTTSPTSQNFVTRSRKALQKFL 1339  
 QY 743 ----- 742  
 Db 1340 PLBETLEKRIIVDVTSTNGKMKHLPFTLTQIDYNEKKGATIQSLADCTSHSI 1399  
 QY 743 ----- 742  
 Db 1400 PQMRSEPLFAKVSFPIRIYILTRVLFQMSHLPASVTKKOSGVSRSHFQAKK 1459  
 QY 743 ----- 742  
 Db 1460 NNLSLAILTMTGQREVGSGTATNSVTYKXVNTVLKPLDPTSGKVELLPVHI 1519  
 QY 743 ----- 742  
 Db 1520 YOKULPFTTSNGSRHLDVSGSLQCTGAIKQNEARNPKVPLRVATESAKTSK 1579  
 QY 743 ----- 742  
 Db 1580 LLDLAMDHNHGTQPKRKNQKSPKTAFKKQTLISLNCSSNHAJALNQK 1639  
 QY 743 ----- 742  
 Db 1640 ELEVWAKQRTKLCSENPVLQKQSSHTTLOSGQRLDDTISYERKCEDFPI 1699  
 QY 743 ----- 742  
 Db 1700 DHDNQSPSPQKTRHTTAAVERLDYQMSSPHYLBNBAQSGSYQFKVVPQFTD 1759  
 QY 743 ----- 742  
 Db 1760 GSFTQPLKGLMELHGLGPIYIARVDNIMTFNQASRPFYSYSLISYEDROQA 1819  
 QY 743 ----- 742  
 Db 1820 EPRANVAPNETTIFMKVQHMAKTKDPCCAKAFSDVDLVDHVSGLGPLLVCIT 1879  
 QY 743 ----- 742  
 Db 1880 NYLPAHQGVTVQKALPFTIDTGSMTFFEMERNCAPQCNQEDTFFENTREHA 1939  
 QY 743 ----- 742  
 Db 1940 INGTIMOTLPEGLMAQDQRYLWLLSGNSHNHSIFSGHVFVTKREYVQALNVP 1999  
 QY 743 ----- 742  
 Db 2000 GVPTVENLPEKAGHREVCILGSHLHAGMSTFLVYSKQQTPLGASQHTRDPTAS 2059  
 QY 743 ----- 742  
 Db 2060 QYQGMAPKLARLYSGSNANSTKPPSKYKVDLLMILHILKTCQAGVSSLYSG 2119  
 QY 743 ----- 742  
 Db 2120 FTMLSGQKQVTSNGSTQTLAVFQWGSQTCQNTNPTTARVTRAPVHPSHS 2179  
 QY 743 ----- 742  
 Db 2180 TIRWELAGCDLNSCMPLGMSKATQDQATBASVFWFVWVWSSKALHILQSRNAR 2239

QY 1327 FQVNSRMLQKQKTKMKTQVTTQVYSLTMSVYVFEPLSSQDQCHQTLFQNKY 1386  
 Db 2240 FQVNSRMLQKQKTKMKTQVTTQVYSLTMSVYVFEPLSSQDQCHQTLFQNKY 2299  
 QY 1387 KYQGNQSFVNSLSDPBLTMYLRHPSQVQVQVQVQVQVQVQVQVQVQVQV 1438  
 Db 2300 KYQGNQSFVNSLSDPBLTMYLRHPSQVQVQVQVQVQVQVQVQVQVQVQV 2351

RESULT 9  
 ; US-10-132-829-4  
 ; Sequence 4, Application US/10132829  
 ; Publication No. US20030044982A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hoechst, Mannheim, R  
 ; TITLE OF INVENTION: Method to treat hemophilia by hepatic gene transfer of Factor V  
 ; TITLE OF INVENTION: with vesicle vector  
 ; CURRENT APPLICATION NUMBER: US/10132, 829  
 ; CURRENT FILING DATE: 2002-04-25  
 ; PRIOR APPLICATION NUMBER: 60/286,314  
 ; NUMBER OF SEQ ID NOS: 2-04-25  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 4  
 ; LENGTH: 2351  
 ; ORGANISM: Homo sapiens  
 ; US-10-132-829-4

Query Match 94.08; Score 7227; DB 15; Length 2351;  
 Best Local Similarity 61.64; Seq No. 6;  
 Matches 1437; Conservative 1; Indels 894; Gaps 1;

QY 1 ATATATGAGNLSHNDMSGLCELPQAFPPRPAFFPNTSVYTKLTFPFTVHFM 60  
 Db 20 ATATATGAGNLSHNDMSGLCELPQAFPPRPAFFPNTSVYTKLTFPFTVHFM 79  
 QY 80 TATATATGAGNLSHNDMSGLCELPQAFPPRPAFFPNTSVYTKLTFPFTVHFM 130  
 Db 121 TATATATGAGNLSHNDMSGLCELPQAFPPRPAFFPNTSVYTKLTFPFTVHFM 139  
 QY 121 TATATATGAGNLSHNDMSGLCELPQAFPPRPAFFPNTSVYTKLTFPFTVHFM 180  
 Db 140 TATATATGAGNLSHNDMSGLCELPQAFPPRPAFFPNTSVYTKLTFPFTVHFM 199  
 QY 181 TATATATGAGNLSHNDMSGLCELPQAFPPRPAFFPNTSVYTKLTFPFTVHFM 240  
 Db 200 TATATATGAGNLSHNDMSGLCELPQAFPPRPAFFPNTSVYTKLTFPFTVHFM 259  
 QY 241 TATATATGAGNLSHNDMSGLCELPQAFPPRPAFFPNTSVYTKLTFPFTVHFM 300  
 Db 260 TATATATGAGNLSHNDMSGLCELPQAFPPRPAFFPNTSVYTKLTFPFTVHFM 319  
 QY 301 TATATATGAGNLSHNDMSGLCELPQAFPPRPAFFPNTSVYTKLTFPFTVHFM 360  
 Db 320 TATATATGAGNLSHNDMSGLCELPQAFPPRPAFFPNTSVYTKLTFPFTVHFM 379  
 QY 361 TATATATGAGNLSHNDMSGLCELPQAFPPRPAFFPNTSVYTKLTFPFTVHFM 420  
 Db 380 TATATATGAGNLSHNDMSGLCELPQAFPPRPAFFPNTSVYTKLTFPFTVHFM 439  
 QY 421 TATATATGAGNLSHNDMSGLCELPQAFPPRPAFFPNTSVYTKLTFPFTVHFM 480  
 Db 440 TATATATGAGNLSHNDMSGLCELPQAFPPRPAFFPNTSVYTKLTFPFTVHFM 499  
 QY 481 TATATATGAGNLSHNDMSGLCELPQAFPPRPAFFPNTSVYTKLTFPFTVHFM 540  
 Db 500 TATATATGAGNLSHNDMSGLCELPQAFPPRPAFFPNTSVYTKLTFPFTVHFM 559  
 QY 541 TATATATGAGNLSHNDMSGLCELPQAFPPRPAFFPNTSVYTKLTFPFTVHFM 600

Db 560 RLASGLIGELLCTYCKEYQKNGKNSKENVILFVSDVMSNSYLTENIQLFNPAG 619  
 Qy 601 VOLQDEPQASINAMISINVTYFDSQLQSLVCLERAVNTYTLISCAQDTFLSVFSGTTPH 660  
 Db 620 VOLQDEPQASINAMISINVTYFDSQLQSLVCLERAVNTYTLISCAQDTFLSVFSGTTPH 679  
 Qy 661 KMYEDTLLTPFSQETVMSNEMGLITLCHSDSPFMRGTPALAKVNSCQKQDVTYE 720  
 Db 680 KMYEDTLLTPFSQETVMSNEMGLITLCHSDSPFMRGTPALAKVNSCQKQDVTYE 739  
 Qy 721 DYVEDI GAVILSLKGNALPEPSF----- 742  
 Db 740 DYVEDI GAVILSLKGNALPEPSF----- 799  
 Qy 743 DYVEDI GAVILSLKGNALPEPSF----- 742  
 Db 740 DYVEDI GAVILSLKGNALPEPSF----- 799  
 Qy 743 DYVEDI GAVILSLKGNALPEPSF----- 742  
 Db 800 XLQVNSSDLLKMLQSPYTHGLSGLSDIQAKVETFSDDPFGCA DSNVSLSEMTFFPQ 859  
 Qy 743 ----- 742  
 Db 860 LHMSGKMWPTFESGLQRLHNEKLGTTATMELKLDKVFSSNTSLISTIPSDNLAAGTN 919  
 Qy 743 ----- 742  
 Db 920 TSLGLPSPMVRHSDQDUTLFGKSGSPFTEGSPSLSEKNSD KLLBSGLMNSQBSN 979  
 Qy 743 ----- 742  
 Db 980 GNVNSSTESGELFKKRAHGPALJATONALFVNSILKTKNTSNTAKRTHIDGSL 1039  
 Qy 743 ----- 742  
 Db 1040 LIENSFVNMQILSDTEBFKVTPLHDBKMLMDKATRLHNSKNTSKSNEMXQVK 1099  
 Qy 743 ----- 742  
 Db 1100 XEGTIPDPADQPMSEFFQGLFELPESARWORTHQKNSLNSGQSPFKQLVSLGPEKSV 1159  
 Qy 743 ----- 742  
 Db 1160 ONFLSEKXNVVKGSEFTKQVGLKEMVPSSENLFLTLNLDLHNTNCKLQBLEK 1219  
 Qy 743 ----- 742  
 Db 1220 KETLQENVLPQIHVVTGTFNFMQILFLSTRQNVGSGDGAVPQDPSRLNDSTNR 1279  
 Qy 743 ----- 742  
 Db 1280 TKCHAFHFKKGBENGLQNGKQVKAATCTRLSPNTSQNVQPSKEMALQFL 1339  
 Qy 743 ----- 742  
 Db 1340 PULETELEKELVDDTSTQNSKNHLPSTLQIDYNEKGAITQPSLQCLTFSSH 1399  
 Qy 743 ----- 742  
 Db 1400 PQANRSLPAKVNSPSPRIPTLYLTVLQUNSHSLPALSFKDQSGQPSHFLQANK 1459  
 Qy 743 ----- 742  
 Db 1460 NULSLALILNTEGQREKVSIGTNSINVTYKVTYLPKQLPKSGKVELLPKVI 1519  
 Qy 743 ----- 742  
 Db 1520 YQKQLPFTESNGSPGHLDDVSGGLQLOTEGAKKNENARPKQVFLVATSEKTSK 1579  
 Qy 743 ----- 742  
 Db 1580 LIQDLDNMHNGTQIPBEKMSQKSPKNTAFKKQDTLISLANCBNAIYANBQKPF 1639  
 Qy 743 -----SNQPLVAKHOREITRTTLQKQREIDYDDTISVENKKEBDIY 786  
 Db 1640 EIEVTAQKQRETELCSNPVLRKHQREIKRTTLQKQREIDYDDTISVENKKEBDIY 1699

Qy 787 DEBNQSPSPQKTRHYPTAAVRLMDYQGMSSPHYLBNQSGSVSPKQVFCBTD 846  
 Db 1700 DEBNQSPSPQKTRHYPTAAVRLMDYQGMSSPHYLBNQSGSVSPKQVFCBTD 1759  
 Qy 847 GSTFOYFAGELNHEHGLGQPTVDAEYDNTQMTTTFENQKSPYSYSLASVEDQKCA 906  
 Db 1760 GSTFOYFAGELNHEHGLGQPTVDAEYDNTQMTTTFENQKSPYSYSLASVEDQKCA 1819  
 Qy 907 ERFKQFVKNENKTYFKMKQDHEMAPTKQBPCKVAMVYEDVLLKQVHSGGLGELVCHT 966  
 Db 1820 ERFKQFVKNENKTYFKMKQDHEMAPTKQBPCKVAMVYEDVLLKQVHSGGLGELVCHT 1879  
 Qy 967 TILNPAKQDQVTOGERALFPTTFDRTKQVTFNEMENQKCAQK ONQDPTVKNVTRFA 1026  
 Db 1880 TILNPAKQDQVTOGERALFPTTFDRTKQVTFNEMENQKCAQK ONQDPTVKNVTRFA 1939  
 Qy 1027 INCVTHQVTRKSAVQKQKTRVLLKMSNENHSHFSGHVTYTKKBYTALNLTPT 1086  
 Db 1940 INCVTHQVTRKSAVQKQKTRVLLKMSNENHSHFSGHVTYTKKBYTALNLTPT 1999  
 Qy 1087 QVETVYMLPSAAGITRHECLLGHBUAQMSTFLVTSNCKQTPLOMASGHRPQITAS 1146  
 Db 2000 QVETVYMLPSAAGITRHECLLGHBUAQMSTFLVTSNCKQTPLOMASGHRPQITAS 2059  
 Qy 1147 QYQGMAPLALNENSGSINWASTPEPSPKLYDPLAHPHITKIQKQAKQKESLYSQ 1206  
 Db 2060 QYQGMAPLALNENSGSINWASTPEPSPKLYDPLAHPHITKIQKQAKQKESLYSQ 2119  
 Qy 1207 FLYMSLDQKQWQTFGNSGTLTAFFQFNVGSSGKKNMFPPIIATRLPHFPHYS 1266  
 Db 2120 FLYMSLDQKQWQTFGNSGTLTAFFQFNVGSSGKKNMFPPIIATRLPHFPHYS 2179  
 Qy 1267 TIRMEKQCDIANSQPLGMSKASDAQITASSYFTFNATWFSKARHLQGRSNMR 1326  
 Db 2180 TIRMEKQCDIANSQPLGMSKASDAQITASSYFTFNATWFSKARHLQGRSNMR 2239  
 Qy 1327 PQVNPXKMLQVQKTKVQVTVQGVKSLTSMYKEPLISSQDQGHQMTLFPQNGV 1386  
 Db 2240 PQVNPXKMLQVQKTKVQVTVQGVKSLTSMYKEPLISSQDQGHQMTLFPQNGV 2299  
 Qy 1387 KYVQCNQDSFTFPVNSLQPLLYRLIRHPSQVPHQALNMEVLGCEAQDLY 1438  
 Db 2300 KYVQCNQDSFTFPVNSLQPLLYRLIRHPSQVPHQALNMEVLGCEAQDLY 2351

## RESULT 10

US-10-172-712-27  
 ; Description: Application US/10172712  
 ; Publication: 1999US012322A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GRIFFIN, JOHN H.  
 ; APPLICANT: GALE, ANDREW J.  
 ; APPLICANT: PELLEGRINI, JEAN-LUC  
 ; APPLICANT: PELLEGRINI, JEAN-LUC  
 ; TITLE OF INVENTION: STABILIZED PROTEINS WITH ENGINEERED DISULFIDE BONDS  
 ; FILE REFERENCE: 4198-400US1  
 ; PRIORITY REFERENCE: 4198-400US1  
 ; CURRENT FILING DATE: 2002-05-17  
 ; PRIOR FILING DATE: 2001-06-14  
 ; PRIORITY REFERENCE: 4198-400US1  
 ; SUPPLEMENTARY INFORMATION: Ver. 2.1  
 ; SEQ ID NO 27  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-172-712-27

Query Match 94.0%; Score 7227; DB 15; Length 2351;  
 Identity 91.6%; Positives 2177; Negatives 134;  
 Matches 1477; Conservative 0; Mismatches 1; Indels 894; Gaps 1;

|    |      |                                                         |      |    |      |                                                          |      |
|----|------|---------------------------------------------------------|------|----|------|----------------------------------------------------------|------|
| Qy | 1    | ATRRVLAUVEANWNSQSDJGELVADAPFVAVPAGSPFPRVAVVTKLIVPVTWLFN | 60   | Db | 1100 | KEOPIPPDQAPFMSFFKPLUPDISAMTQRTKGNKSLMSQOQSPKOLVSLQPKSVBQ | 1159 |
| Db | 20   | ATRRVLAUVEANWNSQSDJGELVADAPFVAVPAGSPFPRVAVVTKLIVPVTWLFN | 79   | Qy | 743  | -----                                                    | 742  |
| Qy | 61   | JAKPRFMMGLGFTTQAKYQVTVTLKAWSPFVSLHANGSPWASGBASQADDTQSO  | 120  | Db | 1160 | QNPJESKQVWVGKEFTQVGLKENVSPSRLPLFNDJLHNHTNKKQKJQRETEK     | 1219 |
| Db | 80   | JAKPRFMMGLGFTTQAKYQVTVTLKAWSPFVSLHANGSPWASGBASQADDTQSO  | 139  | Qy | 743  | -----                                                    | 742  |
| Qy | 121  | REKEDKVPFGGSSHTVYQVLENGPMASDPCLITISLISHVADVLNGLSGALLVCR | 180  | Db | 1220 | KEFLQENVLQVHVTVTKPMKPKPLFSTQKQVGGSDQGVAVPLQDFASJUNSNR    | 1279 |
| Db | 140  | REKEDKVPFGGSSHTVYQVLENGPMASDPCLITISLISHVADVLNGLSGALLVCR | 199  | Qy | 743  | -----                                                    | 742  |
| Qy | 181  | ESLJACEKVTLAKLTLPVDFGKSWHSETQSLMDQDASASAMPDQBTWYWR      | 240  | Db | 1280 | TKGTAHFSKGEENLBSGLHQTKQVKEVACTTISNPTSQNFVTVOSKRALQOFL    | 1339 |
| Db | 200  | ESLJACEKVTLAKLTLPVDFGKSWHSETQSLMDQDASASAMPDQBTWYWR      | 259  | Qy | 743  | -----                                                    | 742  |
| Qy | 241  | SPULGCRHSTVYVYQGTTEVRHSILEGHTFLVNRHQALESPTETZATCLL      | 300  | Db | 1340 | FLBETEKLRIIVDDTSTQNSQNHKLTPSLTQIDVNEKGAITQSPJCSOCTVSHSI  | 1399 |
| Db | 260  | SPULGCRHSTVYVYQGTTEVRHSILEGHTFLVNRHQALESPTETZATCLL      | 319  | Qy | 743  | -----                                                    | 742  |
| Qy | 301  | MDGLQFLCHESHQDQEMAYKYVDSCTREPOLEKMNKEADVDLDSBQVRF       | 360  | Db | 1400 | FOANRSPLFAKVSFPFPIYTVLVPQNHSHPAASVTKKQSCVORSRSHPLQAGK    | 1459 |
| Db | 320  | MDGLQFLCHESHQDQEMAYKYVDSCTREPOLEKMNKEADVDLDSBQVRF       | 379  | Qy | 743  | -----                                                    | 742  |
| Qy | 361  | DDNSESFTQRYVAKQKPKWYHAAEBEDMDYAPLAPDQSYKQVNNKQPGIC      | 420  | Db | 1460 | NNLSLAILLHMTCQDQREVSGISQVATNSVYKVKVTVLKFPLDKTSGKVELLPKVI | 1519 |
| Db | 380  | DDNSESFTQRYVAKQKPKWYHAAEBEDMDYAPLAPDQSYKQVNNKQPGIC      | 439  | Qy | 743  | -----                                                    | 742  |
| Qy | 421  | KYKAPKMAVYDSTFREAQTESGILGFLYQGVGVTLLIPQQAQSRPNVPHGI     | 480  | Db | 1520 | YQKDLPTFTSNGSFOHDLVBSGLLQTCGAKKNENRPGKVFARVATSSSANTPSK   | 1579 |
| Db | 440  | KYKAPKMAVYDSTFREAQTESGILGFLYQGVGVTLLIPQQAQSRPNVPHGI     | 499  | Qy | 743  | -----                                                    | 742  |
| Qy | 481  | TDVRPLYSRLPVYHGLKOPPLPCEIFKYKTVTVYEDGPTKSDPCLTRYYSFVME  | 540  | Db | 1580 | LLOPLANDNHVGTQLPKEHNSQKSPKTAFCCKQDTLSLNACENHAAIAINEQKRP  | 1639 |
| Db | 500  | TDVRPLYSRLPVYHGLKOPPLPCEIFKYKTVTVYEDGPTKSDPCLTRYYSFVME  | 559  | Qy | 743  | -----                                                    | 742  |
| Qy | 541  | RLASGLIPALLCYKESVQGRNOLMSKENVLFSVFDNSESXYLTENIORFLPNAG  | 600  | Db | 1640 | ZIEVWMAQGTBLSQNPVLAKHREBETRTTLAQSQRESDTDTLSVANKCEPDVY    | 1699 |
| Db | 560  | RLASGLIPALLCYKESVQGRNOLMSKENVLFSVFDNSESXYLTENIORFLPNAG  | 619  | Qy | 1700 | DEBQSPSTQKATHTTAAVBLUDVGNSSHSYVJNMQQSPKQVPLVWVFTD        | 1759 |
| Qy | 601  | VOLEDFEPOSNTMSNGVYFDSLQALQVHEVAYVTLISGACTDFLVSFVGSTFH   | 660  | Db | 1760 | DEBQSPSTQKATHTTAAVBLUDVGNSSHSYVJNMQQSPKQVPLVWVFTD        | 1759 |
| Db | 620  | VOLEDFEPOSNTMSNGVYFDSLQALQVHEVAYVTLISGACTDFLVSFVGSTFH   | 679  | Qy | 1820 | EPARVYKNETKTFWVQHMATPOEDFCAMATSDVLEKOVHSGLGPVLVCHT       | 1879 |
| Qy | 661  | KMYVEDTLTPFSSGTVSNMFMHGLTQCHNSDFNMRKATLAKVSCQNDQVTE     | 720  | Db | 1880 | EPARVYKNETKTFWVQHMATPOEDFCAMATSDVLEKOVHSGLGPVLVCHT       | 1939 |
| Db | 680  | KMYVEDTLTPFSSGTVSNMFMHGLTQCHNSDFNMRKATLAKVSCQNDQVTE     | 739  | Qy | 1940 | INGYIMTLPLWMAQDQRLWYLLMSKGNSTHSHFSFSGVTVKKEBIPALVLYP     | 1999 |
| Qy | 721  | DSYVEDISAVLSJONALRPSF-----                              | 742  | Db | 2000 | INGYIMTLPLWMAQDQRLWYLLMSKGNSTHSHFSFSGVTVKKEBIPALVLYP     | 2059 |
| Db | 740  | DSYVEDISAVLSJONALRPSF-----                              | 759  | Qy | 2060 | QYVTEVMSKAGIHWVCELTGBHLKAGNSTLFIYVNSKQTPFLWASCHTPOQTAS   | 2119 |
| Qy | 743  | -----                                                   | 742  | Db | 2120 | QYVTEVMSKAGIHWVCELTGBHLKAGNSTLFIYVNSKQTPFLWASCHTPOQTAS   | 2179 |
| Db | 800  | KIQVSSDOLLKILQSPFHGSLGSLDQANVETFSDFSPCADSNLSJETHFQ      | 859  | Qy | 2180 | QYVTEVMSKAGIHWVCELTGBHLKAGNSTLFIYVNSKQTPFLWASCHTPOQTAS   | 2239 |
| Qy | 743  | -----                                                   | 742  | Db | 2240 | QYVTEVMSKAGIHWVCELTGBHLKAGNSTLFIYVNSKQTPFLWASCHTPOQTAS   | 2299 |
| Db | 860  | LHHSQDMVTFESGLQRLXELKGTAAATEKLODFKVSSTNNLSTPISLQNDGN    | 919  | Qy | 2300 | QYVTEVMSKAGIHWVCELTGBHLKAGNSTLFIYVNSKQTPFLWASCHTPOQTAS   | 2359 |
| Qy | 743  | -----                                                   | 742  | Db | 2360 | QYVTEVMSKAGIHWVCELTGBHLKAGNSTLFIYVNSKQTPFLWASCHTPOQTAS   | 2419 |
| Db | 920  | TSSLGPNVHVDQDLDTLTFCKKSSPLTESGPLSLEENWDSKLESQWSSQSH     | 979  | Qy | 2420 | QYVTEVMSKAGIHWVCELTGBHLKAGNSTLFIYVNSKQTPFLWASCHTPOQTAS   | 2479 |
| Qy | 743  | -----                                                   | 742  | Db | 2480 | QYVTEVMSKAGIHWVCELTGBHLKAGNSTLFIYVNSKQTPFLWASCHTPOQTAS   | 2539 |
| Db | 980  | GNVSSFTSGRLFKYHANGPALLUQNALPVNSILLATKNTSNGSATRKTHIDQSL  | 1039 | Qy | 2540 | QYVTEVMSKAGIHWVCELTGBHLKAGNSTLFIYVNSKQTPFLWASCHTPOQTAS   | 2599 |
| Qy | 743  | -----                                                   | 742  | Db | 2600 | QYVTEVMSKAGIHWVCELTGBHLKAGNSTLFIYVNSKQTPFLWASCHTPOQTAS   | 2659 |
| Db | 1040 | LIENSPSVQHILSDTEFKQVTVLHDMALMDQNALALNNKNTTSSQNMWFOOK    | 1099 | Qy | 2660 | QYVTEVMSKAGIHWVCELTGBHLKAGNSTLFIYVNSKQTPFLWASCHTPOQTAS   | 2719 |
| Qy | 743  | -----                                                   | 742  | Db | 2720 | QYVTEVMSKAGIHWVCELTGBHLKAGNSTLFIYVNSKQTPFLWASCHTPOQTAS   | 2779 |









[illegible]

Db 1819 SKECQPLGASGRTRFOITASQOWAPKALRLHYSGINASTNDPSIKWIDLLA 1878  
 Qy 1184 PMTHIKTCQARQSSSTISQPIWISLDKQNTTSTSTGTUATFFQNVSSGIM 1243  
 Db 1879 PALLHGTQWQARQSSSTISQPIWISLDKQNTTSTSTGTUATFFQNVSSGIM 1938  
 Qy 1244 NTFNPTTIAVILRPHSTASTLWELGQDCUNSCMPLGSKESKASIDAOITASSFT 1303  
 Db 1939 NTFNPTTIAVILRPHSTASTLWELGQDCUNSCMPLGSKESKASIDAOITASSFT 1998  
 Qy 1304 NMFATGSESKALRQGRNARSPQWNPMPKLVDPFKTKVTVTTCQKSLATSY 1363  
 Db 1999 NMFATGSESKALRQGRNARSPQWNPMPKLVDPFKTKVTVTTCQKSLATSY 2058  
 Qy 1364 KEPLTSSGQGHQMTLFFQNKYKQFQNGQDSFFVNSLDPLELTVLRHPQSMWQ 1423  
 Db 2059 KEPLTSSGQGHQMTLFFQNKYKQFQNGQDSFFVNSLDPLELTVLRHPQSMWQ 2118  
 Qy 1434 AURMEVGECAQDLY 1438  
 Db 2119 AURMEVGECAQDLY 2133

## RESULT 15

US-10-187-319-37  
 Publication US/10187319  
 Publication NO. US20030046785A1  
 GENERAL INFORMATION:  
 APPLICANT: Iollar, John S.  
 INVENTOR: Iollar, John S.  
 NUMBER OF SEQUENCES: 1  
 CORRESPONDENCE ADDRESS:  
 ADDRESSES: Greenlee, Wimmer and Sullivan, P.C.  
 STREET: 5370 Manhattan Circle Suite 201  
 CITY: Denver  
 STATE: Colorado  
 COUNTRY: USA

COMPUTER: IBM PC compatible  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 CURRENT NAME: Patent Release #1.0, Version #1.30  
 APPLICATION NUMBER: US/10187,319  
 FILING DATE: 27-Aug-2002  
 CLASSIFICATION: <Unknown>  
 PRIORITY:  
 APPLICATION NUMBER: US 09/523,656  
 FILING DATE: 2000-03-10  
 APPLICATION NUMBER: US 09/037,601  
 FILING DATE: 1997-06-26  
 APPLICATION NUMBER: WO PCT/US97/11155  
 APPLICATION NUMBER: US 08/670,707  
 FILING DATE: 1996-06-26  
 ATTORNEY/INVENTOR:  
 NAME: Greenlee, Lorraine L.  
 REGISTRATION NUMBER: 27,894  
 REFERENCE/DOCKET NUMBER: 75-98K  
 TELEPHONE: 303/499-8080  
 TELEFAX: 303/499-8089

INFORMATION FOR SEQ ID NO: 37:  
 SEQUENCE CHARACTERISTICS:  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SOURCE ORGANISM:  
 SEQUENCE DESCRIPTION: SEQ ID NO: 37:

Query Match 91.04; Score 6227; DB 15; Length 2133;

Best Local Similarity 57.31; Pred. No. 0;  
 Matches 1212; Conservative 106; Mismatches 119; Indels 678; Gaps 59;  
 1 ATRVYGLGAVLEGRSNGQD-LOELVDAAPRPVAKSPFNTSVVYKTLFVETWILF 59  
 Db 2 ALRVYGLGAVLEGRSNGQD-LOELVDAAPRPVAKSPFNTSVVYKTLFVETWILF 79  
 Qy 60 NIAXRPRPMLGLPTIQAEDTVYVITLTKASHPVSLHVAISVWASGASDQDTS 119  
 Db 80 SVAREPAPLGLPTIQAEDTVYVITLTKASHPVSLHVAISVWASGASDQDTS 139  
 Qy 120 QREKDDKVPFGSSHTVWQVTKAGNPASDPLCTLTSTLSKDVKDSGLTGALLVC 179  
 Db 140 QREKDDKVPFGSSHTVWQVTKAGNPASDPLCTLTSTLSKDVKDSGLTGALLVC 199  
 Qy 180 QREKAKETQVLTHTLLFANVDFGKSMISTEISLQKSDASARAPKQHTVNGV 239  
 Db 200 QREKAKETQVLTHTLLFANVDFGKSMISTEISLQKSDASARAPKQHTVNGV 259  
 Qy 240 RSLPGLGCHKASVYVHVIQNGTPEVHSIFLEHTFLVNRHQSLSEISPTLTATL 299  
 Db 260 RSLPGLGCHKASVYVHVIQNGTPEVHSIFLEHTFLVNRHQSLSEISPTLTATL 319  
 Qy 360 LMDLQQLPCHSSYHQDQMBAYVVDSCPEPQLMNRHREABYDDIDTDSRDVVR 359  
 Db 320 LMDLQQLPCHSSYHQDQMBAYVVDSCPEPQLMNRHREABYDDIDTDSRDVVR 378  
 Qy 360 PDDNSPSTFIEISVAKKPKVWYVTAAREEDMDYAPLPDUSKYQVLANSPRT 419  
 Db 379 LQGDVGFPIQHSVAKKPKVWYVTAAREEDMDYAPLPDUSKYQVLANSPRT 438  
 Qy 420 GRXYKQVPMALVTDEFTKTEALQHSGLGLPLGVGQDGLLIIPKQASDPATVNG 479  
 Db 439 GRXYKQVPMALVTDEFTKTEALQHSGLGLPLGVGQDGLLIIPKQASDPATVNG 498  
 Qy 480 ITDVRPIKXRLKQVHLDKQPLTLOEIEFKVWVTEQDPTSDPCLTYSSTFNM 539  
 Db 499 ITDVRPIKXRLKQVHLDKQPLTLOEIEFKVWVTEQDPTSDPCLTYSSTFNM 558  
 Qy 540 ERLASGLQMLTKQVSKQKQKQNMKNVNLVSVNENRMYLTENLQPLVPA 599  
 Db 559 ERLASGLQMLTKQVSKQKQKQNMKNVNLVSVNENRMYLTENLQPLVPA 618  
 Qy 600 QVQEDPESQNTMNSITVPSILQVSLCHETVATLISGAQDTPSFFSGITFK 659  
 Db 619 QVQEDPESQNTMNSITVPSILQVSLCHETVATLISGAQDTPSFFSGITFK 678  
 Qy 660 HMYVEDTLTLPFSSTVFMNENGLMICHNSDFNKGKALLKVSQCDQDGY 719  
 Db 679 HMYVEDTLTLPFSSTVFMNENGLMICHNSDFNKGKALLKVSQCDQDGY 738  
 Qy 720 EDSYEDSVILSKMNAIEPFSFQ-  
 Db 739 DAYTELPGGLSKGVLEFRAPKSPSSAQQQTPTSDVEDVDPQSGERTQAL 798  
 Qy 746 ----- 745  
 Db 799 ELSVPSGDSGMLQVAPRESSDQLQARHEDDYLPARENTPAARALPELH 858  
 Qy 746 ----- 745  
 Db 859 HSARVTFPEKELKEDKQSSDLKSTPTTISDLSAERTHSLGHPQVNER 918  
 Qy 746 ----- 745  
 Db 919 SQGVAVGRKNSHFGVPLGSTEEDHESLGSNVSVESDGIKEKRAKGLATND 978  
 Qy 746 ----- 745  
 Db 979 DVYFKVNLVTKNKAFTYKTRKTHIDQALLTERASATFMDQYVAGLVNMI 1038  
 Qy 746 ----- 745

Search completed: December 9, 2003, 17:09:44  
Job time : 80 secs

|      |                                                           |      |      |      |                                                           |
|------|-----------------------------------------------------------|------|------|------|-----------------------------------------------------------|
| 1039 | XBFJLGNPLJSEHCPSSPELLTSSGSGKVGSGQGRJRWAVREELXKNGKWNLSF    | 1099 | 746  | 1099 | XBFJLGNPLJSEHCPSSPELLTSSGSGKVGSGQGRJRWAVREELXKNGKWNLSF    |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |
| 1099 | YIFLTLNSADVQNUHSHSQKSEBMBREBLQVEKQYDLPVYATGTATQKPLNLFHQST | 1158 | 745  | 1099 | YIFLTLNSADVQNUHSHSQKSEBMBREBLQVEKQYDLPVYATGTATQKPLNLFHQST |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |
| 1158 | BSYVEFGDGSHPAPQDSGSSJNDSABRTHLHAHFSALREEPALQVQWTCPPSNAV   | 1218 | 745  | 1158 | BSYVEFGDGSHPAPQDSGSSJNDSABRTHLHAHFSALREEPALQVQWTCPPSNAV   |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |
| 1218 | PRBYQSLKQILPLLEEIXPERGVOTVATSTWSSSESPQLQAGKNNLSLPTILEKAGQ | 1278 | 745  | 1218 | PRBYQSLKQILPLLEEIXPERGVOTVATSTWSSSESPQLQAGKNNLSLPTILEKAGQ |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |
| 1278 | QKLSLALCKSSAGPLGKLEKAVLSSAGLSEASQJAEFLKJVWRELDLPQTSWISC   | 1338 | 745  | 1278 | QKLSLALCKSSAGPLGKLEKAVLSSAGLSEASQJAEFLKJVWRELDLPQTSWISC   |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |
| 1338 | AGUDLQOEIPLQKTEGFWLAKVNRPORTSKLLQCPHWKNEHSLEKPSYKALATDII  | 1398 | 746  | 1338 | AGUDLQOEIPLQKTEGFWLAKVNRPORTSKLLQCPHWKNEHSLEKPSYKALATDII  |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |
| 1398 | EJLPLRHSHSHTANNOQOEITQJEDAMTQKQFOLCARFPLKJLQDILQISUPTOP   | 1458 | 763  | 1398 | EJLPLRHSHSHTANNOQOEITQJEDAMTQKQFOLCARFPLKJLQDILQISUPTOP   |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |
| 1458 | QOEBIIDYDITTSVEMKUEDFQITDDEGSPFQKCTHNTFIJAAVEALWDYSGSSSHV | 823  | 763  | 1458 | QOEBIIDYDITTSVEMKUEDFQITDDEGSPFQKCTHNTFIJAAVEALWDYSGSSSHV |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |
| 823  | EDKNOJDTFSTGSGHJQYDILGDEBNOQPHSPQCTHNTFIJAAVEALWDYSGSSSHV | 1518 | 823  | 823  | EDKNOJDTFSTGSGHJQYDILGDEBNOQPHSPQCTHNTFIJAAVEALWDYSGSSSHV |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |
| 1518 | LEKQJGQVFKKVPKFPADGSPQSTGSLGMLHJLGLPITRAVEDINMVFEN        | 1578 | 883  | 1518 | LEKQJGQVFKKVPKFPADGSPQSTGSLGMLHJLGLPITRAVEDINMVFEN        |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |
| 1578 | CASRPSYFSSLSIYEDQAGKPRKFWKNEKTYPKWQENAPKQFOPCKKWAY        | 1638 | 943  | 1578 | CASRPSYFSSLSIYEDQAGKPRKFWKNEKTYPKWQENAPKQFOPCKKWAY        |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |
| 1638 | PSVDLQKQVSGSLGILLCRANTLNAAGQVOTQVEPALFTIPQSTSWTFENVR      | 1698 | 1003 | 1638 | PSVDLQKQVSGSLGILLCRANTLNAAGQVOTQVEPALFTIPQSTSWTFENVR      |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |
| 1698 | KCPACHPQWEDPQKENTHRYNGVMTQPLVMAQOQRTWLTSLKSGNSHNSTH       | 1758 | 1063 | 1698 | KCPACHPQWEDPQKENTHRYNGVMTQPLVMAQOQRTWLTSLKSGNSHNSTH       |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |
| 1758 | PSGHVFTYKPEEYKALNTYVPGVETVHSPKSVKQVIMBELLQHLQAGSTPLVY     | 1818 | 1123 | 1758 | PSGHVFTYKPEEYKALNTYVPGVETVHSPKSVKQVIMBELLQHLQAGSTPLVY     |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |
| 1818 | SNKQCPGLMAQSHQDPTQASQOYGMCPKLALHFGSSNANWCTSPFPMIUVOLA     | 1878 | 1183 | 1818 | SNKQCPGLMAQSHQDPTQASQOYGMCPKLALHFGSSNANWCTSPFPMIUVOLA     |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |
| 1878 | PMHGHITMGQAGKSSSIYQEIITVSLDGNWQSGNSTGTQVAVFQVDSGLCH       | 1938 | 1243 | 1878 | PMHGHITMGQAGKSSSIYQEIITVSLDGNWQSGNSTGTQVAVFQVDSGLCH       |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |
| 1938 | NIENPPIYANVILRHPHVSJSTLJAMELQJLSCSLPQWESLXASDQATNASFT     | 1303 | 1303 | 1938 | NIENPPIYANVILRHPHVSJSTLJAMELQJLSCSLPQWESLXASDQATNASFT     |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |
| 1303 | NWATPSSVABHJLQENABRQKQWJLQKQWQKQWQKQWQKQWQKQWQKQWQK       | 1363 | 1363 | 1303 | NWATPSSVABHJLQENABRQKQWJLQKQWQKQWQKQWQKQWQKQWQKQWQK       |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |
| 1363 | NIPATPSSVABHJLQENABRQKQWJLQKQWQKQWQKQWQKQWQKQWQKQWQK      | 1423 | 2058 | 1363 | NIPATPSSVABHJLQENABRQKQWJLQKQWQKQWQKQWQKQWQKQWQKQWQK      |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |
| 1423 | KBEFLJSSQSGHQTFFPKWQKQWQKQWQKQWQKQWQKQWQKQWQKQWQKQWQK     | 1483 | 2118 | 1423 | KBEFLJSSQSGHQTFFPKWQKQWQKQWQKQWQKQWQKQWQKQWQKQWQKQWQK     |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |
| 2118 | ALAEVACEQDLY1438                                          |      |      | 2118 | ALAEVACEQDLY1438                                          |
| Qb   |                                                           | Qb   |      | Qb   |                                                           |

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2003, 17:02:15 / Search time 28 seconds  
1938.949 Million cell updates/sec

Title: US-10-047-257-1

Partial score: 1  
Sequence: 1 ATERYLGVNELSHYQMSD.....WRIQALNEVLGCEADLY 1438

Scoring table: BIOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database : PIR\_76.\*

1: p121.\*

2: p122.\*

3: p123.\*

4: p124.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result | No.    | Score | Query Match | Length | DB ID  | Description             |
|--------|--------|-------|-------------|--------|--------|-------------------------|
| 1      | 7227   | 94.0  | 2351        | 1      | B2HU   | coagulation factor      |
| 2      | 6227   | 81.0  | 2133        | 2      | 742763 | coagulation factor      |
| 3      | 6192   | 80.5  | 2313        | 2      | A47004 | coagulation factor      |
| 4      | 23272  | 30.8  | 2211        | 1      | KF805  | coagulation factor      |
| 5      | 23269  | 30.8  | 2183        | 2      | 742764 | coagulation factor      |
| 6      | 1731.5 | 22.5  | 1069        | 1      | K0HU   | ferroxidase EC 1.11.1.1 |
| 7      | 1104   | 14.4  | 216         | 2      | A42268 | agg protein precursor   |
| 8      | 663    | 8.6   | 427         | 2      | Q4935  | agg protein precursor   |
| 9      | 657    | 8.5   | 463         | 1      | A36479 | milk fat globule m      |
| 10     | 443    | 8.5   | 407         | 2      | J00948 | PF47 protein - pig      |
| 11     | 432    | 8.5   | 407         | 2      | S51741 | PAS-77 protein anti     |
| 12     | 635    | 8.3   | 427         | 2      | S74211 | coagulation factor      |
| 13     | 543    | 7.1   | 869         | 2      | A25945 | AS antigen precursor    |
| 14     | 443    | 7.1   | 977         | 1      | J00948 | milk fat globule p      |
| 15     | 443    | 7.1   | 977         | 1      | S51741 | adipocyte transcri      |
| 16     | 306.5  | 4.0   | 3113        | 2      | 852083 | transcription repr      |
| 17     | 262    | 3.4   | 845         | 2      | Q52586 | neurexin IV - fruit     |
| 18     | 221    | 2.9   | 719         | 2      | 852739 | neurexin IV - fruit     |
| 19     | 207.5  | 2.7   | 717         | 2      | T31349 | hypothetical prote      |
| 20     | 206.5  | 2.7   | 737         | 2      | T15615 | receptor tyrosine       |
| 21     | 193    | 2.5   | 913         | 2      | A48280 | protein-tyrosine k      |
| 22     | 185.2  | 2.5   | 876         | 2      | A49508 | protein-tyrosine k      |
| 23     | 185.2  | 2.4   | 910         | 2      | A53137 | tyro 10 receptor k      |
| 24     | 173.5  | 2.3   | 819         | 2      | T48859 | paranodin - rat         |
| 25     | 171    | 2.2   | 1381        | 2      | T31083 |                         |

## RESULT 1

Species: Homo sapiens (man)

C/Species: Homo sapiens (man)

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

R/Gitachier, J.; Wood, W.I.

Hum. Mol. Genet. 1, 199-200, 1992

Abstract: Sequence of the exon-containing regions of the human factor VIII gene.

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I

Accession: I54318, A00525, I580597 A215847, A26177, A123487, A03986, I635577, S66445, I



Db 860 LHSIGDHPFTBSGLQRLNEKLGTTATTEALKLDFKVSSTNNLSTTTPDNLAAAGTN 919  
 Qy 743 ----- 742  
 Db 920 TSLSGPPSPVHVQDQLDTLLFGKKSSPLTBSGQPLSLSEENDNKLBSGLANSQESH 979  
 Qy 743 ----- 742  
 Db 980 QHNVSTESGRLFKGDKGAPALLTKONALPKVLSILAKTKNTSNGATNKKTHLOSL 1039  
 Qy 743 ----- 742  
 Db 1040 LIENSPSVQNLISLDESREKFTPLIDHMLDKDNATALLNUNSNKSTSSKNRQKQK 1099  
 Qy 743 ----- 742  
 Db 1100 KEGTIPFDQNPNSFFKALFLPSASMIQTHQKNSLSQSGQPPQLVSLGPKSVYEG 1159  
 Qy 743 ----- 742  
 Db 1160 QNPLSEKNNVVKGEFTVGLGVKVPVPSGNNFYNLMDLHNNHNRQKKEKTEK 1219  
 Qy 743 ----- 742  
 Db 1220 KHTLQENVVLPQHTYGTGNPKHRLPLASTQNVGSGTDGMVPLQDPSLNDSTNR 1279  
 Qy 743 ----- 742  
 Db 1280 TKGTATFSKXGKEENBELGNQTKYKIACTRTISNPTSQQNFVQSRKALQPL 1339  
 Qy 743 ----- 742  
 Db 1340 PLRETELEKRIIVDTSTQNSKNKHLTPSTLTQIDNTEKEKATQSPSLDCLTSHSI 1399  
 Qy 743 ----- 742  
 Db 1400 PQRNSPLPIAKYSFSTIRPTLYTRVLPDNGSHLPAASYKCKSCQVBSHFLOALK 1459  
 Qy 743 ----- 742  
 Db 1460 NNLISLAILMTGQREVGSGTSAIVSYKTKVNTVLPKDPKTSCKVELLPVHI 1519  
 Qy 743 ----- 742  
 Db 1520 YQKDLFTPTBSNGSLGLDLVGGSLQQTGGAIKNNEARPGKVPFLVATSSANTPSK 1579  
 Qy 743 ----- 742  
 Db 1580 LLDPLADNHNHGTQPKKEKQSKPKPTAKKZDTLISLAKNSHATAINHSQNP 1639  
 Qy 743 ----- 742  
 Db 1640 ELEWNAKQGTBELCSNPPPLVQKQEBTNTTLQSLKELNADIVDTLSVNRGEPHII 1699  
 Qy 787 DDENQSPSPQCKTHRTITATVRLMYQVSSSHPHVLNAGSVQPKVAVVQEPFD 846  
 Db 1700 DDEQDPSSTQKTHRTITATVRLMYQVSSSHPHVLNAGSVQPKVAVVQEPFD 1759  
 Qy 847 GSTPTQYFGLNHLGLGPTTPAEDVDNNTVTFNQASPSYSLSLSTVEDQCA 906  
 Db 1760 GSTPTQYFGLNHLGLGPTTPAEDVDNNTVTFNQASPSYSLSLSTVEDQCA 1819  
 Qy 907 EPRKQFAPKNTIKYKQHQHAPFDQPCANAFVSDVLEQVHSLGAPLVCHT 966  
 Db 1820 EPRKQFAPKNTIKYKQHQHAPFDQPCANAFVSDVLEQVHSLGAPLVCHT 1879  
 Qy 967 NTLNPAIGQVTVQSEALFTTFIDTSYVTFNEMERNCRAPONQEDPTKENTRPA 1026  
 Db 1880 NTLNPAIGQVTVQSEALFTTFIDTSYVTFNEMERNCRAPONQEDPTKENTRPA 1939  
 Qy 1027 INGYIMTLPGVAQDORLRWLLKSGNSNTHSIFSHSVTFVTKKRYEYALNLYP 1086  
 Db 1940 INGYIMTLPGVAQDORLRWLLKSGNSNTHSIFSHSVTFVTKKRYEYALNLYP 1999

Qy 1087 GYVETVBLPLKAGHVRVCELGHLHAGNSTPLVYVSNKQCTPLGASGRPOITAS 1146  
 Db 2000 GYVETVBLPLKAGHVRVCELGHLHAGNSTPLVYVSNKQCTPLGASGRPOITAS 2059  
 Qy 1147 QYGVQWAPKALHVRVCELGHLHAGNSTPSPKSYKDLLAPMLHGTGKQCARQKFSYLSQ 1206  
 Db 2060 QYGVQWAPKALHVRVCELGHLHAGNSTPSPKSYKDLLAPMLHGTGKQCARQKFSYLSQ 2119  
 Qy 1207 PLTWSLDCKQWOTGNSGTGLTAVPQNVSSCKNENFPPILARYILRHPHTYSYR 1266  
 Db 2120 PLTWSLDCKQWOTGNSGTGLTAVPQNVSSCKNENFPPILARYILRHPHTYSYR 2179  
 Qy 1267 PLTWSLDCKQWOTGNSGTGLTAVPQNVSSCKNENFPPILARYILRHPHTYSYR 1326  
 Db 2180 PLTWSLDCKQWOTGNSGTGLTAVPQNVSSCKNENFPPILARYILRHPHTYSYR 2239  
 Qy 1327 PQQNVKQWQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 1386  
 Db 2240 PQQNVKQWQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQ 2299  
 Qy 1387 KYVQNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSG 1438  
 Db 2300 KYVQNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSG 2351

## RESULT 2

T42763

coagulation factor VIII precursor - pig

C1Species: Sus scrofa domestica (domestic pig)

C1Date: -Jan-2000 Sequence\_revision 11-Jan-2000 #ext\_change 09-Jun-2000

C1Name: F142763

R1Gollard, P.

submitted to the EMBL Data Library, August 1996

A1Reference number: Z22269

A1Accession number: F142763

A1Status: preliminary; translated from GB/EMBL/DBJ

A1Molecule type: mRNA

A1Residues: 1-2133 &lt;L05&gt;

A1Superfamily: coagulation factor VIII

A1Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilias A; pl

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

F142763/Domain: signal sequence factor VIII

Query Match 81.0%; Score 6207; DB 2; Length 2133;

Best Local Similarity 57.3%; Pred No. 0;

Matches 1212; Conservative 106; Mismatches 119; Indels 678; Gaps 3;

Qy 1 ATKNSYQVKEGKSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSG 59

Db 2 ATKNSYQVKEGKSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSG 79

Qy 60 MIAKPRPAMGLGPTTPAEDVDNNTVTFNQASPSYSLSLSTVEDQCA 119

Db 80 SVAREPAMGLGPTTPAEDVDNNTVTFNQASPSYSLSLSTVEDQCA 139

Qy 120 QREKDDYVFGGSHYVQVLEKNGPASPCLCTYSLSHVQLKQANSLGALLVC 179

Db 140 QREKDDYVFGGSHYVQVLEKNGPASPCLCTYSLSHVQLKQANSLGALLVC 199

Qy 180 RESLAKNQLQKTLKTLVPAEDVCKSHGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSG 239

Db 200 RESLAKNQLQKTLKTLVPAEDVCKSHGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSGNSG 259

Qy 240 RSLPLGICGKHSYVHYVQVLEKNGPASPCLCTYSLSHVQLKQANSLGALLVC 299

Db 260 RSLPLGICGKHSYVHYVQVLEKNGPASPCLCTYSLSHVQLKQANSLGALLVC 319

Qy 300 LKDLQGLPCLCHSHQSDQMEATVYVSDVCEPCLAKMKNNEAEADYDIDDLTDSKQVYR 359















```

QY 741 SPSQWVYK.....751
Db 1099 LVGSVPAEHSPTTPAQDPQTHSTTDRSYRSPFELSQGLDYDLSDIFVPDDIGLTSFF 1158
QY 752 751
Db 1159 PQSQKSSFSDDQALFSSDLSPITSPELDQTIIVPLDQLLLSPEDNQKSTSPDLQ 1218
QY 752 751
Db 1219 VPLSPDNKNTSSPDQLQVSLSPDNQKSSPDQLQVSLSPDNQKNTSPDLQVLSPD 1278
QY 752 EITRTLSQDE.....767
Db 1279 DNQMTSPDLQVPLSDNQKNTSSPDQLQVLPREDNQKNTLUSGLPLASDQKST 1338
QY 768 767
Db 1339 DLUTLSPPDQGVTLSPDLQQLPLSDNSQVTVSPDLGLTLSPDNFELLANPOLQVTL 1398
QY 768 767
Db 1399 PDLITQWPAHNGHAGSADPQASPPFDSQASLPELNLWTLHPLDUTLHPPSPPL 1458
QY 768 ISVMEKKEDDIDVDENSDNSRSPQ 798
Db 1459 NPTLSKRNVLVWGLRGVDDGVTVSPSPERDIDEDYDPAWTVNPDTRDTHV 1518
QY 799 KKRHR.....YFAVRLWLDYGMSSSPHNRAGSQSVQ---FK 837
Db 1519 NSGNPPTTANLYLGHGQHKFYFLIAEHLWNTAFQSEM--DHEDTGHPTDITTK 1576
QY 838 KVPQETFGSTFQCLYRGLNELHGLGLPIAEVADENIMVTFRQASRPVSFSLIS 897
Db 1577 KVVRKTLGDTFTSPDPAEYSEHLGITGVTAETVDVDTQVTPNLASRPSYAHGLS 1636
QY 898 YE-----EDCGSGERKFNQZKNTETPKVQCHMAPKDECKAMVPSVDVL 949
Db 1637 YKSSBEKQYEDSPFQEDDANVQNSSTVYVHATSGSPNGSGACAWYATVANY 1696
QY 950 EKQVHSGELGLVCHTLNGLNANGVTVQVFALPTFLPETSRYFTFENBKRSPC 1009
Db 1697 ERLHSGGLGLVLCRGTGLNEMERTPNMRGEPVLLPVPDEKSMYTKS--KGR-- 1751
QY 1010 MWQDPTKPTVHAGVINYVQITPQLVAGDQGLVTLWLSNGSNVTLWLSHGW 1069
Db 1752 -RISPEKNAKATYALNGLTYR-LPLAMFQEWLHLLAMGSRDHWVHFQTL 1808
QY 1070 TVRKSEYVQVALYVQVETVQMLPKAGLVIVZCLGHHVLAQSTFPAVYKQST 1129
Db 1809 LDNRKTHQGLQVWLLASFTLSEAKSKGMLLDVEVAGVAGTFFLITDECKN 1868
QY 1130 PKMASGHTIRPDTAGQVQGVQVGLRAGLRYKSNVYKTE-----PFS-NYVLLA 1183
Db 1869 PMLGSLTSSDLSLASELITLWFLALNLAGSNVMSIEXTALDLPFKMLQVDMQK 1928
QY 1184 PMIHITKTCQASQFSSLYLQPTITMSLDKQKQVTRTMSVGLVYFPWDSGIGH 1243
Db 1929 EVWVGLQGLQALSKLCTFEQVAYSSDQVWQVLFAGSGSWVYFGNSDSTIK 1988
QY 1244 NTMPDPIIARTVIRPLHPSYVISTIRMLZMLCDLNSCHPLGSKASISDAQTASYFT 1303
Db 1989 NLLPPLIATLHNPSTNPLMLLQGLVGLVSTPLFLEDEGLQVQVLSNPK 2048
QY 1304 NNPAT-NSPSABRLQGSNMPQVNNKMLQVQVQVQVQVQVQVQVQVQVQV 1362
Db 2049 SHWQYWFSLDGLNGLVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 2108
QY 1363 VEEFLISSGQCHQHTLFQNKV--KVFQNGQSTFFVMSLDPLTLRYLHQPQVY 1420
Db 2109 VVSYLSQVSDQVAMKTPKQSGWQVQVQVQVQVQVQVQVQVQVQVQVQV 2168
QY 1421 HQTALREMLGCEAQDLY 1438

```

Db 2169 QSTALREMLGCE--DTY 2183

# RESULT 7

ferroxidase (EC 1.16.3.1) precursor [validated] - human

NAlternate names: ceruloplasmin

NContains: ferroxidase long form (CP-1); ferroxidase short form (CP-2)

NDate: 31-Aug-1998 Sequence revision 12-May-1995 #text\_change 08-Dec-2000

NAccession: A25443; A24155; A35450; A00524; I59067

NKochinsky, M.L.; Funk, W.D.; van Oost, B.A.; MacGillivray, R.T.A.

NProc. Natl. Acad. Sci. U.S.A. 83: 186-190, 1986

NProc. Natl. Acad. Sci. U.S.A. 83: 186-190, 1986

NReference number: A25443; MUID:8625973; PMID:2873574

NAccession: A25443

NMolecule type: mRNA

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443

NAccession: A25443







F:1-21/Domain: signal sequence Metastat predicted <SIG>

F:28-60/Domain: EGF homology <EGF>

F:168-107/Domain: discoidin I amino-terminal homology <DNI>

F:110-267/Domain: discoidin I amino-terminal homology <DNI>

F:1271-427/Domain: discoidin I amino-terminal homology <DNI>

Query Match

Best Local Similarity 40.3%; Pred. No. 1.6e-36;

Matches 142; Conservative 61; Mismatches 123; Indels 26; Gaps 6;

QY 1105 SCLECHLGMSTLTVYSNK-----CQPLMASGHIDPDTASQY-- 1149

DB 78 KCLVETDQGG--DFITFYTCQPVSGYCHCECTSLGEGIAISQISNVYMG 135

QY 1150 ----CGAPKRLARLVSGSNW--STKEPFSIKVLLAPMLHIGTCGRKFSSLY 1203

DB 136 PMLQGRVGFELARLTQIVNATVASTSYDSKVFQVFLRQVRVSTQGRAGRAY 195

QY 1204 LSOFTIMWGLDQKQVCGENCTGLAVFGNSKQENKFNIPDIARYLPHRTHYS 1263

DB 196 LKTFVAVVAGLGRFFQDSQDQKPFMDQNNLSLKNVNPFLTAQYLRVTVSCH 255

QY 1264 IKTQVTELWGLDQKQVCGENCTGLAVFGNSKQENKFNIPDIARYLPHRTHYS 1320

DB 256 RGCTLRFELLCGELCEPELAKNTPDQSTLASSYKYNLAFVPHLGRDLNQG 315

QY 1321 KSNARVQVQKFKHQVQKQVNTQVQKQVNTQVQKQVNTQVQKQVNTQVQKQV 1380

DB 316 KINATVSSAKEMQLVDQVQKQVNTQVQKQVNTQVQKQVNTQVQKQVNTQVQK 375

QY 1381 PQNGKVTQVQKQVNTQVQKQVNTQVQKQVNTQVQKQVNTQVQKQVNTQVQK 1432

DB 376 EQGQTSVATVQNLQNNSHKNTFEKFPKRVAVVPLVSHNRTITLVELLGC 427

Result 11

A36473

milk fat globule membrane protein - mouse

C:Species: Mus musculus (house mouse)

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #ext\_change 10-Sep-1999

R:Stubbs, J.D.; Lekusis, C.; Singer, K.L.; Bai, A.; Yuzaki, D.; Srinivasan, U.; Parry, C.

Proc. Natl. Acad. Sci. U.S.A. 87, 8417-8421, 1990

Article: cDNA cloning of a mouse mammary epithelial cell surface protein reveals the ex

Accession: A36473

A:Status: preliminary

A:Molecule type: mRNA

A:Cross-references: GB:M8337; NID:g19142; PID:AA39534.1; PID:g19143

C:Keywords: membrane protein

F:28-60/Domain: EGF homology <EGF>

F:147-303/Domain: discoidin I amino-terminal homology <DNI>

F:307-463/Domain: discoidin I amino-terminal homology <DNI>

Query Match

Best Local Similarity 42.2%; Pred. No. 4.7e-36;

Matches 135; Conservative 63; Mismatches 110; Indels 12; Gaps 5;

QY 1124 SNVQCTVQNSGHIDPDTASQY-----CGAPKRLARLVSGSNW--STKEPFS 1175

DB 145 ASKSCVTCGALALSOISVSYVPMGFMQGFELARLVPTGLVNVHNSDKSE 204

QY 1176 KIVQDLARVTHIGTCGRKFSSLYISQVTHIMVQLVQVQKQVNTQVQKQVNTQVQK 1235

DB 205 KIVQDLARVTHIGTCGRKFSSLYISQVTHIMVQLVQVQKQVNTQVQKQVNTQVQK 263

QY 1236 VNSGSGIKFNIPDIARYLPHRTHYSIKVQLVQVQKQVNTQVQKQVNTQVQKQV 1295

DB 264 LNNNSLVNNTVLTQYTRLVVPSCHQCTLCELLELQACLEPELKNVTPDQ 323

QY 1296 ITASSYFT--NWPA-TNPSKARLVHQSNAVRQVNNIKEMQLVDQKPMKVQVTTQ 1352

DB 324 MSASSYKTYNLAFAVPHLGRDLNQGKINAKTASNSAKEMQLVDQKQVQVQVITQ 383

QY 1353 QKSLTQKYNKVEELSSQCHQHTLPQNKVKYQKQVQDSFTVPSVLDPLPALT 1412

DB 384 GQHTQVSTVKYVNSDQVQVTEECSSQVFGQILQNNSHKNTFEKFPKRVAV 443

QY 1413 RHPQGRVQKQVNTQVQKQVNTQVQKQVNTQVQKQVNTQVQKQVNTQVQK 1432

DB 444 KVLVPSVSHNRTITLVELLGC 463

Result 12

T11743

p847 protein - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 16-Jun-1999 #sequence\_revision 16-Jul-1999 #ext\_change 08-Sep-2002

R:Breslin, M.; Vogel, T.; Calvete, J.J.; Thole, H.R.; Schmidtke, J.; Matsuda, T.; Toep

Biol. Reprod. 59, 1057-1064, 1998

Article: Molecular cloning and characterization of p84, a novel boar sperm-associated

Accession: T11743

A:Status: preliminary; translated from GB/E05L/DB0J

A:Molecule type: mRNA

A:Cross-references: B04L:Y11683; NID:g4552927; PID:CAA72379.1; PID:g4552928

A:Function:

A:Description: may be involved in membrane remodeling and/or function as a sperm cellu

F:6-40/Domain: EGF homology <EGF>

Query Match

Best Local Similarity 38.5%; Pred. No. 1.6e-36;

Matches 139; Conservative 62; Mismatches 125; Indels 28; Gaps 6;

QY 1105 ECLIGLHGMSTLTVYSNK-----CQPLMASGHIDPDTASQY-- 1149

DB 58 EKVVDVING--DVFTFYCKPGVYGHICILNQLNQTALADQLSASSHLG 115

QY 1150 ----CGAPKRLARLVSGSNW--STKEPFSIKVLLAPMLHIGTCGRKFSSLY 1203

DB 116 PMLQGRVGFELARLTQIVNATVASTSYDSKVFQVFLRQVRVSTQGRAGRAY 195

QY 1204 LSOFTIMWGLDQKQVCGENCTGLAVFGNSKQENKFNIPDIARYLPHRTHYS 1263

DB 176 KCLVETDQGG--DFITFYTCQPVSGYCHCECTSLGEGIAISQISNVYMG 235

QY 1264 IKTQVTELWGLDQKQVCGENCTGLAVFGNSKQENKFNIPDIARYLPHRTHYS 1320

DB 236 RGCTLRFELLCGELCEPELAKNTPDQSTLASSYKYNLAFVPHLGRDLNQG 295

QY 1321 KSNARVQVQKFKHQVQKQVNTQVQKQVNTQVQKQVNTQVQKQVNTQVQKQV 1380

DB 296 KFWATVSSAKEMQLVDQVQKQVNTQVQKQVNTQVQKQVNTQVQKQVNTQVQK 355

QY 1381 PQNGKVTQVQKQVNTQVQKQVNTQVQKQVNTQVQKQVNTQVQKQVNTQVQK 1432

DB 356 EQGQTSVATVQNLQNNSHKNTFEKFPKRVAVVPLVSHNRTITLVELLGC 409

Result 13

S65138

cytoproterin antigen MG95/53

C:Species: Bos primigenius taurus (cattle)

C:Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #ext\_change 07-Aug-1998

Accession: S65138; G48394

Article: Molecular cloning of cytoproterin antigen MG95/53 recognized by monoclonal I

Biochim. Biophys. Acta 1245, 88-91, 1995

Article: Molecular cloning of cytoproterin antigen MG95/53 recognized by monoclonal I



Search completed: December 9, 2003, 17:08:39  
Job time : 56 secs

| Result | No.    | Score | Query Match | Length | DB          | ID     | Description |
|--------|--------|-------|-------------|--------|-------------|--------|-------------|
| 1      | 7227   | 94.0  | 2351        | 1      | F88_HUMAN   | P00451 | homo sapien |
| 2      | 4192   | 80.5  | 2119        | 1      | F88_MOUSE   | O06159 | mus mus mus |
| 3      | 4192   | 80.5  | 2119        | 1      | F88_MOUSE   | O06159 | mus mus mus |
| 4      | 2389.5 | 31.1  | 2224        | 1      | FAS_HUMAN   | P12259 | homo sapien |
| 5      | 2372   | 30.8  | 2211        | 1      | FAS_BOVIN   | Q83107 | bos taurus  |
| 6      | 1737.5 | 25.5  | 1665        | 1      | CRU_HUMAN   | P08214 | homo sapien |
| 7      | 1737.5 | 25.5  | 1665        | 1      | CRU_HUMAN   | P08214 | homo sapien |
| 8      | 1684.5 | 21.9  | 1059        | 1      | GERU_MOUSE  | P61435 | rattus norv |
| 9      | 1683.5 | 21.3  | 1062        | 1      | GERU_MOUSE  | P61435 | rattus norv |
| 10     | 597    | 8.5   | 467         | 1      | MFEM_MOUSE  | P70390 | rattus norv |
| 11     | 597    | 8.5   | 467         | 1      | MFEM_MOUSE  | P70390 | rattus norv |
| 12     | 650    | 8.5   | 409         | 1      | MFEM_PIG    | P79385 | mus mus mus |
| 13     | 635    | 8.1   | 427         | 1      | MFEM_BOVIN  | Q85114 | bos taurus  |
| 14     | 635    | 8.1   | 427         | 1      | MFEM_BOVIN  | Q85114 | bos taurus  |
| 15     | 459.8  | 6.1   | 937         | 1      | NP21_HUMAN  | Q08421 | homo sapien |
| 16     | 464.5  | 6.0   | 925         | 1      | NP22_MOUSE  | O35376 | mus mus mus |
| 17     | 462.5  | 6.0   | 931         | 1      | NP22_MOUSE  | O35376 | mus mus mus |
| 18     | 459.5  | 5.8   | 914         | 1      | NP21_CHECK  | P79795 | gallus gall |
| 19     | 459.5  | 5.8   | 914         | 1      | NP21_CHECK  | P79795 | gallus gall |
| 20     | 446.5  | 5.1   | 923         | 1      | NP21_MOUSE  | P07333 | mus mus mus |
| 21     | 443    | 5.8   | 928         | 1      | NP21_XENTLA | P26824 | xenopus lae |
| 22     | 362.5  | 4.0   | 323         | 1      | NP21_HUMAN  | O14786 | homo sapien |
| 23     | 362.5  | 4.0   | 323         | 1      | NP21_HUMAN  | O14786 | homo sapien |
| 24     | 266    | 3.5   | 764         | 1      | CPX2_MOUSE  | Q84215 | mus mus mus |
| 25     | 261    | 3.4   | 756         | 1      | CPX2_MOUSE  | Q84215 | mus mus mus |
| 26     | 261    | 3.4   | 756         | 1      | CPX2_MOUSE  | Q84215 | mus mus mus |
| 27     | 200.5  | 3.4   | 280         | 1      | XLR1_FURCU  | Q06475 | hugu rubrip |
| 28     | 200.5  | 3.4   | 280         | 1      | XLR1_FURCU  | Q06475 | hugu rubrip |
| 29     | 223    | 3.2   | 224         | 1      | XLR1_HUMAN  | O15537 | mus mus mus |
| 30     | 223    | 3.2   | 224         | 1      | XLR1_HUMAN  | O15537 | mus mus mus |
| 31     | 217    | 2.8   | 722         | 1      | CP2M_MOUSE  | Q92100 | mus mus mus |
| 32     | 217    | 2.8   | 722         | 1      | CP2M_MOUSE  | Q92100 | mus mus mus |
| 33     | 211    | 2.7   | 1264        | 1      | NRX4_MOUSE  | Q06633 | homo sapien |
| 34     | 211    | 2.7   | 1264        | 1      | NRX4_MOUSE  | Q06633 | homo sapien |
| 35     | 208    | 2.7   | 1310        | 1      | CTM4_MOUSE  | Q59547 | mus mus mus |

- EX MEDLINE=91093266; PubMed=1989735;  
 RA Meyer A. van Schijndel S.B., Michrs C., Huttnar M.B., Verbeet M.P.,  
 RA "Salifation of Tyr1680 of human blood coagulation factor VIII is  
 RT essential for the interaction of factor VIII with von Willebrand  
 RT factor". *Biochem. J.* 266:740-746 (1991).  
 RN [7]  
 RN SUPPATTION.  
 RP MEDLINE=92207952; PubMed=1554716;  
 RA "Identification and functional importance of tyrosine sulfate  
 RT residues within recombinant factor VIII.";  
 RN [18]  
 RN SUPPATTION BY NMR OF 2322-2343.  
 RP MEDLINE=95200924; PubMed=7893714;  
 RA Gilbert G.E., Baleja J.D.; from the CD domain of factor VIII forms an  
 RT "Membrane-binding peptide" that is determined by NMR spectroscopy.";  
 RN [19]  
 RN REVIEW ON MOLECULAR BASIS OF HEMA.  
 RA Gieschehr J.; PubMed=1592642;  
 RA "The molecular basis of hemophilia A.";  
 RN [20]  
 RN ANN. N.Y. ACAD. SCI. 614:89-96 (1991).  
 RP REVIEW ON MOLECULAR BASIS OF HEMA.  
 RA MEDLINE=89088506; PubMed=2491949;  
 RA White G.C. II, Shoemaker C.B.;  
 RT "Factor VIII gene and hemophilia A.";  
 RN [21]  
 RN REVIEW ON MOLECULAR BASIS OF HEMA.  
 RA MEDLINE=95245332; PubMed=7726145;  
 RA "Molecular biology of factor VIII deficiency in hemophilia A.";  
 RN [22]  
 RN HUM. MUTAT. 5:1-23 (1995).  
 RP MEDLINE=94235048; PubMed=1012775;  
 RA Gieschehr J., Wood M.J., Shuman M.A., Larn R.M.;  
 RT "Identification of a missense mutation in the factor VIII gene of a  
 RN Sicilian hemophilic patient.";  
 RN [23]  
 RP VARIANT HEMA PRO-2135; PubMed=3120181;  
 RA MEDLINE=88096340; PubMed=2810040;  
 RA Phillips J.A. III, Janco R.L., Hoyer L.W.;  
 RT "A novel missense mutation in the factor VIII gene identified by  
 RN analysis of amplified hemophilia DNA sequences.";  
 RN Nucleic Acids Res. 15:9797-9805 (1987).  
 RP VARIANT HEMA GLN-2228.  
 RA MEDLINE=88191889; PubMed=2833855;  
 RA Youseoufian H., Antonarakis S.E., Bell W., Griffin A.M.;  
 RT "Missense and missense mutations in hemophilia A: estimate of the  
 RN relative mutation rate at CG dinucleotides.";  
 RN Am. J. Hum. Genet. 42:718-725 (1988).  
 RP VARIANT HEMA GLY-291.  
 RA MEDLINE=88220354; PubMed=28355904;  
 RA Youseoufian H., Wong C., Acronis S., Platokoukis H., Kazanian H.H. Jr.,  
 RA Antonarakis S.E.; hemophilia A resulting from Glu--Gly substitution  
 RT in exon 7 of the factor VIII gene.";  
 RN [16]  
 RN HUM. GENET. 42:867-871 (1988).  
 RP VARIANT HEMA CYS-1708.  
 RA MEDLINE=89274593; PubMed=2499363;  
 RA O'Brien D.P., Tuddenham E.G.;  
 RT "Purification and characterization of factor VIII 1,489-Cys, a  
 RA functional cofactor occurring in a patient with severe hemophilia  
 RN A.";  
 RN Blood 73:2117-2122 (1989).  
 RP MEDLINE=90001543; PubMed=2506948;  
 RA Shima M., Ware J., Yoshioka A., Fukui H., Fulcher C.A.;  
 RT "An arginine to cysteine amino acid substitution at a critical  
 RN site in the heavy chain of a dysfunctional factor VIII molecule.";  
 RN [18]  
 RN Blood 74:1542-1517 (1989).  
 RP VARIANT HEMA LEU-189.  
 RA MEDLINE=90057687; PubMed=2510835;  
 RA "A novel missense mutation in exon 4 of the factor VIII:C gene  
 RT resulting in moderately severe hemophilia A.";  
 RN [20]  
 RN Blood 74:2688-2691 (1989).  
 RP VARIANT HEMA LEU-2326.  
 RA MEDLINE=89197216; PubMed=2495245;  
 RA Inaba H., Fujimaki M., Kazanian H.H. Jr., Antonarakis S.E.;  
 RT "Wild hemophilic A resulting from Arg-to-Leu substitution in exon 26  
 RN of the factor VIII gene.";  
 RN Hum. Genet. 81:335-338 (1989).  
 RP VARIANT HEMA HIS-391.  
 RA MEDLINE=89088506; PubMed=2499892;  
 RA Araki M., Inaba H., Higuchi M., Antonarakis S.E., Kazanian H.H. Jr.,  
 RA Fujimaki M., Hoyer L.W.;  
 RT "Direct characterization of factor VIII in plasma: detection of a  
 RN mutation affecting the thrombin cleavage site  
 RN Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281 (1989).  
 RP VARIANT HEMA CYS-1708.  
 RA MEDLINE=95245332; PubMed=2104765;  
 RA Araki M., Higuchi M., Antonarakis S.E., Kazanian H.H. Jr.,  
 RA Phillips J.A. III, Janco R.L., Hoyer L.W.;  
 RT "Characterization of a thrombin cleavage site mutation (Arg 1689 to  
 RN Cys) in the factor VIII gene of a hemophilic patient with  
 RN a plasma-reacting inhibitor.";  
 RN Blood 75:384-389 (1990).  
 RP VARIANT HEMA GLN-2208 AND LEU-2326.  
 RA MEDLINE=90123183; PubMed=2105106;  
 RA Casula L., Murru S., Pecorara M., Riscaldi M.S., Restagno G.,  
 RA Mancuso G., Morfini M., de Biasi R., Baudo P., Carbonara A.;  
 RT "VII gene mutation and thrombin cleavage site mutation in the factor  
 RN Blood 75:662-670 (1990).  
 RP VARIANT HEMA CYS-391.  
 RA MEDLINE=90123183; PubMed=2105106;  
 RA Patterson J.K., Mowley J.B., Tuddenham E.G.;  
 RT "CmH. hemophilia A due to a missense mutation (372--Cys) at the  
 RN internal heavy chain thrombin cleavage site.";  
 RN [24]  
 RN Blood 75:652-670 (1990).  
 RP VARIANT HEMA PHE-1659 AND CYS-1708.  
 RA MEDLINE=90123183; PubMed=2105106;  
 RA Higuchi M., Wong C., Kochan L., Olek K., Acronis S., Kasper C.K.,  
 RT "Characterization of mutations in the factor VIII gene by direct  
 RN sequencing of amplified genomic DNA.";  
 RN [16]  
 RN Genomics 6:65-71 (1990).  
 RP VARIANT HEMA CYS-1728 AND ASP-1941.  
 RA MEDLINE=90169988; PubMed=2106480;  
 RA Trosman M.D., Higuchi M., Kasper C.K., Antonarakis S.E.,  
 RA Kazanian H.H.;  
 RT "Using gradient gel electrophoresis to detect point  
 RN mutations in the factor VIII gene.";

| Query Match           | 94.0% ; Score 7227 ; DB 1 ; Length 2351 ;                       |                                                                   |
|-----------------------|-----------------------------------------------------------------|-------------------------------------------------------------------|
| Best Local Similarity | 61.6% ; Pred. No. 0 ;                                           |                                                                   |
| Matches 137 ;         | Conservative 0 ;                                                | Matches 1 ; Indels 894 ; Gaps 1 ;                                 |
| Qy                    | 1 ATRRYGLGAVELSHQVQSDGLGFLVDARPPRVKPSPPNTVVLKTLVFTVFLFN 60      | 1040 LTRNGSVQMIIEEDTEKKVTLTHORALADRATALLRNINSNTVTSKRMVQOK 1099    |
| Db                    | 20 ATRYLLGVAVELSHQVQSDGLGFLVDARPPRVKPSPPNTVVLKTLVFTVFLFN 79     | 743 ----- 742                                                     |
| Qy                    | 61 IAKRPFMGLLGFTIQASVYDVVTVLLKMAHSHVSLHANGSVTKASBQAYDQDSQ 120   | 1100 KESFPDQAPQNSFFNMLFPLSARWIOHQHNSLNSQGPSPQVLSLGPENSVBG 1159    |
| Db                    | 61 IAKRPFMGLLGFTIQASVYDVVTVLLKMAHSHVSLHANGSVTKASBQAYDQDSQ 139   | 743 ----- 742                                                     |
| Qy                    | 121 REKEDAVVFGGSHYVQVVLKGVKPPASDPLCTVSLSHVQLXOLNSGLGALVC 180    | 1160 QNFLSENKVVVKGCEFTKQVGLKEMVFPSSNLFNLNLDLHNNNTNQEKKIOEIEK 1219 |
| Db                    | 140 REKEDAVVFGGSHYVQVVLKGVKPPASDPLCTVSLSHVQLXOLNSGLGALVC 199    | 743 ----- 742                                                     |
| Qy                    | 181 EGSIAKELQTLKTLTALVAVFDGSKSHSETKSLAQDRAASAPAPKQVTVYNR 240    | 1220 KETLQENWLPQLHVTVGTGNFKMLFLSLSTQVSSVYQAVAPVADQFSLNDSYNE 1279  |
| Db                    | 200 EGSIAKELQTLKTLTALVAVFDGSKSHSETKSLAQDRAASAPAPKQVTVYNR 259    | 743 ----- 742                                                     |
| Qy                    | 241 SLPGLQCHRSYVTVYVQMGTPVSHVTLFEGHTVAVNRHQAASLFTFTFAQLL 300    | 1280 TKCHTAHPSKGGEENLGLGNQTKVETKACTTTSPTSQGVTVQSRKALQFRL 1339     |
| Db                    | 260 SLPGLQCHRSYVTVYVQMGTPVSHVTLFEGHTVAVNRHQAASLFTFTFAQLL 319    | 743 ----- 742                                                     |
| Qy                    | 301 MDGLQPLFCHISQHDQAYVYKVCQCEPQAPKQNEARVYQDQATSGKQVPR 360      | 1340 PLRETELEXLIVDQSTQNSKMLFSTLTLDINKEKALQTQSLADCLUTSISIT 1399    |
| Db                    | 320 MDGLQPLFCHISQHDQAYVYKVCQCEPQAPKQNEARVYQDQATSGKQVPR 379      | 743 ----- 742                                                     |
| Qy                    | 361 DDNGSESTQITRQVAKHVTQVTAABRQVAPLPLAPRPSVQVYKQANRQPLG 420     | 1400 PQANRPLPIANKVSFSPITPYTVTVFQNSHSLPAASVTKDQSGVQBSHPLOAKX 1459  |
| Db                    | 380 DDNGSESTQITRQVAKHVTQVTAABRQVAPLPLAPRPSVQVYKQANRQPLG 439     | 743 ----- 742                                                     |
| Qy                    | 421 RYKAKENAVTVFVPRVAVLQHRGSLQPLAVLQVGTLLITPKQASPVMTIFPH 480    | 1460 NKLSIALILBERTQVQVSGVSGTASVTVYKVVNTVLPKEDPLSPCKVELLPVHI 1519  |
| Db                    | 440 RYKAKENAVTVFVPRVAVLQHRGSLQPLAVLQVGTLLITPKQASPVMTIFPH 499    | 743 ----- 742                                                     |
| Qy                    | 481 TDVPLPSLPGVYALDPLLPGLFICKQTVVDEQSTQSPCLTRYTSYSPVME 540      | 1520 YOKDLFTPTTNSGSHLDVGGSLQOTEGA IKMGNARPKGVFVRYATPSSANTPSK 1579 |
| Db                    | 500 TDVPLPSLPGVYALDPLLPGLFICKQTVVDEQSTQSPCLTRYTSYSPVME 559      | 743 ----- 742                                                     |
| Qy                    | 541 ADLASGLQPLLCYKESVQDQATMDGNVILFVTVENNSVLYTNIOTPLNPAG 600     | 1580 LLOPLANDVNTQITPKRWSQKSPKTAFFKXDTILSNACENHATAINEQOKP 1639     |
| Db                    | 560 ADLASGLQPLLCYKESVQDQATMDGNVILFVTVENNSVLYTNIOTPLNPAG 619     | 743 ----- 742                                                     |
| Qy                    | 601 VQLEDFPOANRMSINQVTFDGLQVCLVAVYVTLAQCTDVLAVFPGSTPH 660       | 1640 RIEVTKAAQKTEBLCSNPFFVAKHQBETLTKTLOQOQETIDVDTISVHKQDQVY 786   |
| Db                    | 620 VQLEDFPOANRMSINQVTFDGLQVCLVAVYVTLAQCTDVLAVFPGSTPH 679       | 787 DEBQDQKQPTBLTANVTRVHVNQSVSSHWVAKQAGQVQVYKQVYVFTD 846          |
| Qy                    | 661 KMYEDTLPLFPFTVFSVPMQGLMIGCHNSDFNRQNTALLVSCDKNTDYIE 720      | 1700 DEBQDQKQPTBLTANVTRVHVNQSVSSHWVAKQAGQVQVYKQVYVFTD 1759        |
| Db                    | 680 KMYEDTLPLFPFTVFSVPMQGLMIGCHNSDFNRQNTALLVSCDKNTDYIE 739      | 847 GSVTPVYKQKSNHGLQVFTVAKSDONINVTNQLSHPTSYSSLISTEDQCA 906        |
| Qy                    | 721 DSYEDISALLSNNALBSEF----- 742                                | 1760 GSVTPVYKQKSNHGLQVFTVAKSDONINVTNQLSHPTSYSSLISTEDQCA 1819      |
| Db                    | 740 DSYEDISALLSNNALBSEF----- 742                                | 907 EPRNPKVQKNTVTKVQVHNNLPTDQEPQANVTVDEQVSGSLGFLVCHT 966          |
| Qy                    | 743 ----- 742                                                   | 1820 EPRNPKVQKNTVTKVQVHNNLPTDQEPQANVTVDEQVSGSLGFLVCHT 1879        |
| Db                    | 800 KIONVSSDLMLKQSPHPLGLSLDQKAVETSDQSPQADSNNSISEMTHFQ 859       | 967 MTLNPAQNVTVFALFTTIDSTKSYFTENNEKAPCMQMDPTKENTREFA 1026         |
| Qy                    | 743 ----- 742                                                   | 1880 MTLNPAQNVTVFALFTTIDSTKSYFTENNEKAPCMQMDPTKENTREFA 1939        |
| Db                    | 860 LHSQDGVTFESGLQRLNEKLGCTYATLKLQFVKSSTNSNLTSTPDSNLANGTN 919   | 1027 INCTYMDTLPLVMAQDQRTWLLWLGNSNINISHFSGVFTVKKERYKALVLP 1086     |
| Qy                    | 743 ----- 742                                                   | 1940 INCTYMDTLPLVMAQDQRTWLLWLGNSNINISHFSGVFTVKKERYKALVLP 1999     |
| Db                    | 920 TSSLSPSPVHDSQDLDTLLFQKESSTVTSQGPLSLSEBNNDSKLLGSLMWSBSSW 979 | 1087 QVTFVPMPSKAGIHWVCELLGHJHAGNCTFLFVYSNKQTPLOWASHGRDQFTAS 1146  |
| Qy                    | 743 ----- 742                                                   | 2000 QVTFVPMPSKAGIHWVCELLGHJHAGNCTFLFVYSNKQTPLOWASHGRDQFTAS 2059  |
| Db                    | 980 GKNVSTESGLFKKGAHPALLTNKALFVNSILAKTNKTSNNSATNRKHIDQSL 1039   | 1147 QYQVAPKPLARLHYSQINASTKEPFNKVDLAPMLIHQIKTQANQKPSLSIQ 1206     |
| Qy                    | 743 ----- 742                                                   | 2060 QYQVAPKPLARLHYSQINASTKEPFNKVDLAPMLIHQIKTQANQKPSLSIQ 2119     |
| Db                    |                                                                 | 1207 PTMYSLDKQKQTVGNSGTGLWYFPGVDSGGEKNIENPPIIARYILFHPHYSIRS 1266  |

















FT DOMAIN 1748 1890 PLASTOCYANIN-LIKE 6.  
 FT DOMAIN 1894 2048 F5/6 TYPE C 1.  
 FT DOMAIN 2053 2208 F5/6 TYPE C 2.  
 FT SITE 1034 1305 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).  
 FT SITE 1564 1635 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).  
 FT DISULFID 167 193 PROBABLE.  
 FT DISULFID 192 225 PROBABLE.  
 FT DISULFID 732 1325 PROBABLE.  
 FT DISULFID 1894 2208 BY SIMILARITY.  
 FT MOD\_RES 697 597 SULFATION (POTENTIAL).  
 FT MOD\_RES 730 730 SULFATION (POTENTIAL).  
 FT MOD\_RES 1513 1513 SULFATION (POTENTIAL).  
 FT MOD\_RES 1529 1529 SULFATION (POTENTIAL).  
 FT MOD\_RES 1541 1541 SULFATION (POTENTIAL).  
 FT MOD\_RES 225 225 N-LINKED (GLNAC).  
 FT CARBOHYD 239 239 N-LINKED (GLNAC).  
 FT CARBOHYD 297 297 N-LINKED (GLNAC).  
 FT CARBOHYD 327 327 N-LINKED (GLNAC).  
 FT CARBOHYD 460 460 N-LINKED (GLNAC).  
 FT CARBOHYD 553 553 N-LINKED (GLNAC).  
 FT CARBOHYD 587 587 N-LINKED (GLNAC).  
 FT CARBOHYD 756 756 N-LINKED (GLNAC).  
 FT CARBOHYD 774 774 N-LINKED (GLNAC).  
 FT CARBOHYD 780 780 N-LINKED (GLNAC).  
 FT CARBOHYD 922 922 N-LINKED (GLNAC).  
 FT CARBOHYD 964 964 N-LINKED (GLNAC).  
 FT CARBOHYD 1044 1044 N-LINKED (GLNAC).  
 FT CARBOHYD 1053 1053 N-LINKED (GLNAC).  
 FT CARBOHYD 1071 1071 N-LINKED (GLNAC).  
 FT CARBOHYD 1078 1078 N-LINKED (GLNAC).  
 FT CARBOHYD 1094 1094 N-LINKED (GLNAC).  
 FT CARBOHYD 1490 1490 N-LINKED (GLNAC).  
 FT CARBOHYD 1550 1550 N-LINKED (GLNAC).  
 FT CARBOHYD 1690 1690 N-LINKED (GLNAC).  
 FT CARBOHYD 1997 1997 N-LINKED (GLNAC).  
 FT CARBOHYD 2196 2196 N-LINKED (GLNAC).  
 FT VARIANT 587 592 NTPA -> T (IN VARIANT 2).  
 SQ SEQUENCE 2211 AA, 248981 NF; CBBF908738667C5 CMC64.  
 Query Match. 30.88; Score 2172; DB 1; Length 2211;  
 Best Local Similarity 26.21; Pred. No. 8.8e-147;  
 Matches 594; Conservative 265; Mismatches 487; Indels 924; Gaps 35;  
 QY 3 RYVLGAVELSDYQWSDGLCPVMDAPRRPVKPFEPFETVYKTLVFTVHLNLTIA 62  
 DB 32 RQYVAAQGIWNRV-----PESHV-----SKPFET-----FKAVIREYAY-FOKE 76  
 QY 63 KPPPPVGLLOPTQAEVDVTVTLTANWASHVPLSHANGVSKASCAVDQDSORE 122  
 DB 77 KQSTGLSGGLPTLAEVADIMKPKHAKKPLHAGIAGIKKESFGASDHLPLME 136  
 QY 123 KEDDQKPGSTGKQVQWGLKNGADPGLCLTVKYSLSHLDWANGKJLWANGK 192  
 DB 137 KADDAVAPQRYTYEIMTISEHSGFTDTPCLTHYTVSLVADVFNGLGLPLCKAG 196  
 QY 183 SLAKENQTL--HEVTLFVPSGSHMSFTNGLMDORDASAMAPRQVWYVVR 240  
 DB 197 TLTEGQVQKQKQVLMHAFVDSKVNCTNS-----LATVNGVYNG 240  
 QY 241 SIPLQICRHYSYVHYVGMGTTPVFEHITFLEHSTFLVNRQASLEISPTFTVQTL 300  
 DB 211 MDQITVCAHRIKSHLLGGLGELPFLHFNQGLQVQKHKIENATLVNASTLVNKKY 300  
 QY 301 TMDQGLFLFCHISSHQDGMAYVYVDSCTEEFQLRMKNNEEADYDDDLTDSBNQVRF 360  
 DB 301 SPFGKMTIASLIPHFQGMQVAVDIDKCKCKKTPK-----LTRO----- 343  
 QY 361 DQNGSFQFQSVKAKKPKYKWKVIAEEDQVAVLPADPQKSYQVYANGQRIQ 420  
 DB 344 -----FHLKRWETFAEVEVDVATIPANMKYSLKLFONGNRIG 388  
 QY 421 RYKQKPMATDITFTR--EALQHSGLLPGLYQVGTLLIPFQKASRPVYTH 478  
 DB 389 KTKKVVYVQVQDSPTKELEDPSSEGLGILIEAQLVUTLKVFNWASRVSIVTH 448  
 QY 479 GIT-----DVRPLSRPQVGVKLOFPILPEFKIATVTVGQPTKSPCL 529  
 DB 449 QTFSPVQWNSSTSSCNTANVR-----KQTYTWNILSDPTFENQAQL 500  
 QY 530 TTYTSPVQWNSSTSSCNTANVR-----KQTYTWNILSDPTFENQAQL 500  
 DB 501 TTYTSPVQWNSSTSSCNTANVR-----KQTYTWNILSDPTFENQAQL 500  
 QY 590 NIQRLFNAPQVLEDFEFOANIME-----SINGYVFSIQ-LSVLRVAVYLSIG 643  
 DB 561 MYTKENPKYKADDPAPFENSNMNTFALNGTVPSTPIQCPDQVQKCSVG 620  
 QY 644 AQDTSEVPSGYTFKHHQVYEDVTLTFPESGTVFMSNPLQWILACQDQVPRQMT 703  
 DB 621 TQNDLTHPTFHSFYCKHEDVTLTFPQOSVTVYVQWVMTLVNNSFKRLR 680  
 QY 704 ALKVVSCDKQVQ-----YVDS-----VEDISVLL----- 731  
 DB 681 LRFQDQCTNRDDDSYRIYFPGSGTAMTKKHDSSTIEDNDAEDQDEALILGL 740  
 QY 732 -----SKNWAETPESQ----- 744  
 DB 741 RFRNSSLQCKDEMLTALAEKDSGFPPFSANSLDSSSRSHVRLAKRFAELK 800  
 QY 745 -----NPPVLK----- 750  
 DB 801 TLHLLEPAMCSLEHAGLDRKSNALNPPAHESSFYSDPDHPLSDVGVSLPFGT 860  
 QY 751 -----RQK-----ETVTTLASQ-----BET 766  
 DB 861 FKKRPAKQKPVQVQGAQAKHKSCTFPAKHTTLLQKNSSSSRMGWEDIPSDLL 920  
 QY 769 ----- 768  
 DB 921 LQKQDPYKILGEMHVLSEKSGYI1QDANKNTVNLANSFONDSRTWGENIPFNSHG 980  
 QY 769 ----- 768  
 DB 981 KQSGHPTLVTRKPLQQRDRNSLKSLGLPLTRKKEKEKPAVHPLSPRSHLR 1040  
 QY 769 -----DYVD----- 772  
 DB 1041 GVNASFSRDRNSHLLPASNYSLSIDLQTFPSNLSLAASLPHQDTPNDVTSQT 1100  
 QY 773 -----TISVWCKEEDFYD-----BNQS----- 794  
 DB 1101 SSPDLPTVPSFENQVFTFQDSPTHTTASRNSPDPTHTTAPRNSFPQPSQIF 1160  
 QY 795 ----- 794  
 DB 1161 NYDLAKNALTPTQVQPSFPLELWQVATSDLSQPSISFDLQGMALSPQOQBSLFD 1220  
 QY 795 ----- 794  
 DB 1221 QVTSLSFPLSQBSLSFDLQVTLSPDPSQBSLFDLQVTLSPDPSQBSLFDLQVTL 1280  
 QY 795 ----- 794  
 DB 1281 PDQOBSLSFDLQVTLSPDPSQBSLFDLQVTLSPDPSQBSLFDLQVTLSPDPSQ 1340  
 QY 795 ----- 794

[illegible]

|    |        |      |      |     |             |                                                               |      |                                        |                     |              |
|----|--------|------|------|-----|-------------|---------------------------------------------------------------|------|----------------------------------------|---------------------|--------------|
| FT | REPEAT | 1285 | 1293 | 14. | FT          | CARBHYD                                                       | 1073 | 1073                                   | N-LINKED (GLCNAc..) | (POTENTIAL). |
| FT | REPEAT | 1294 | 1302 | 15. | FT          | CARBHYD                                                       | 1089 | 1089                                   | N-LINKED (GLCNAc..) | (POTENTIAL). |
| FT | REPEAT | 1303 | 1311 | 16. | FT          | CARBHYD                                                       | 1174 | 1174                                   | N-LINKED (GLCNAc..) | (POTENTIAL). |
| FT | REPEAT | 1311 | 1319 | 17. | FT          | CARBHYD                                                       | 1480 | 1480                                   | N-LINKED (GLCNAc..) | (POTENTIAL). |
| FT | REPEAT | 1331 | 1338 | 18. | FT          | CARBHYD                                                       | 1490 | 1490                                   | N-LINKED (GLCNAc..) | (POTENTIAL). |
| FT | REPEAT | 1339 | 1347 | 20. | FT          | CARBHYD                                                       | 1597 | 1597                                   | N-LINKED (GLCNAc..) | (POTENTIAL). |
| FT | REPEAT | 1339 | 1347 | 20. | FT          | CARBHYD                                                       | 1737 | 1737                                   | N-LINKED (GLCNAc..) | (POTENTIAL). |
| FT | REPEAT | 1357 | 1365 | 21. | FT          | CARBHYD                                                       | 1886 | 1886                                   | N-LINKED (GLCNAc..) | (POTENTIAL). |
| FT | REPEAT | 1357 | 1365 | 21. | FT          | CARBHYD                                                       | 1896 | 1896                                   | N-LINKED (GLCNAc..) | (POTENTIAL). |
| FT | REPEAT | 1356 | 1374 | 23. | FT          | CARBHYD                                                       | 2243 | 2243                                   | N-LINKED (GLCNAc..) | (POTENTIAL). |
| FT | REPEAT | 1375 | 1383 | 24. | FT          | CARBHYD                                                       | 2243 | 2243                                   | N-LINKED (GLCNAc..) | (POTENTIAL). |
| FT | REPEAT | 1384 | 1392 | 25. | SQ          | SEQUENCE                                                      | 2258 | AA; 256078 MW; 915989E0076A2ACC CRC64; |                     |              |
| FT | REPEAT | 1384 | 1392 | 25. | Query Match |                                                               |      |                                        |                     |              |
| FT | REPEAT | 1402 | 1410 | 27. | Similarity  | 30.74; Score 2357.5; Db 1; Length 2258;                       |      |                                        |                     |              |
| FT | REPEAT | 1411 | 1419 | 28. | Match       | 587; Conservative 278; Mismatches 485; Indels 963; Gaps 33;   |      |                                        |                     |              |
| FT | REPEAT | 1420 | 1428 | 29. | Match       | 587; Conservative 278; Mismatches 485; Indels 963; Gaps 33;   |      |                                        |                     |              |
| FT | REPEAT | 1438 | 1447 | 30. | QY          | 3 RRYVLGAVLMDWMSDLGELPVDNRPFRVFPPFNTSVVYKTLFVETVHLFNIA 62     |      |                                        |                     |              |
| FT | REPEAT | 1447 | 1455 | 31. | Db          | 32 QYVYVAQLSINHTFE-----PHTSSSPFATB--FKVIVHRYEAY-FOKE 76       |      |                                        |                     |              |
| FT | REPEAT | 1456 | 1464 | 33. | QY          | 63 KRPFWNGLGPTQTAETVTVITLKNWASHVSLVAVGVSKASGAEVDDQSGRE 122    |      |                                        |                     |              |
| FT | REPEAT | 1465 | 1473 | 34. | Db          | 77 RPSMSGLSLGPTLADGVQDKWHPNCKGSLHPQGLKIKSGFASGAPHTFIVE 136    |      |                                        |                     |              |
| FT | REPEAT | 1483 | 1491 | 36. | QY          | 123 KEDDKVFGSGHTVQVLEKNGMSPADPLCTVSYSLVDVLNGLSGLLACKCBG 182   |      |                                        |                     |              |
| FT | REPEAT | 1492 | 1500 | 37. | Db          | 137 RODVAVPTEHTTENNLSDESGPHNDPCLCTVSYSLVDVLNGLSGLLACKCBG 196  |      |                                        |                     |              |
| FT | REPEAT | 1501 | 1509 | 38. | QY          | 183 SLAKCTQTL--HEFTLFAVFDGKSHSTSTSLNLAQDRAASANAPDHTVGVYVR 240 |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | Db          | 197 TLREGIQMFKQVTPVAVPVDSSKMQSSS-----LATTGVGVNG 240           |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | QY          | 241 ELGLICGRKSVYVHVLGWTTPRVASIFLECHTFVLRHQASLEISPTPLTAQTL 300 |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | Db          | 241 TMDVTCVMDIISHLGLSGHSGVSLFSHFQSVLQGHKVKSAITLVSTNTATVY 300  |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | QY          | 301 MDLQGLFLCHSSHQDHGNEAVYVQSCPPEQLMNQNEEASDYDDIDTDSMDVVR 360 |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | Db          | 301 EPGQKWFISLPHFQAGQNYVDIKQCAKTRPK-----LTRD----- 343         |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | QY          | 361 DDNSPFIQLSIVAKSHKPTVHVIIAEEDNDVAVLADORSVAYQVANGDQIG 420   |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | Db          | 344 -----REHIGRWFTLAEEVINDVATFIPANQKYSALFNDFNQIG 388          |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | QY          | 421 KYVCEVPMYATDTEPR--BALQWESGLGLPQVGVQGLIITFQNASRNPVYH 478   |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | Db          | 389 KHKVAKYVQQSFTFLRPNPNKNDKILGVTRAQVRLTAVFNQNASRNPVYH 448    |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | QY          | 479 GIT-----PRPFLKRGKGVKHLKOPFLAKHFNKVVTVYVDFQKSGDQRY 532     |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | Db          | 449 VTFSPEDVNSSSDNNNTNR-----AVQGEYTTKWLLESDPNDQAQLTP 503      |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | QY          | 533 YSGVFNWDLASLGLPILCYKESVQKQNKMSHREHFLSVDFRWRVLENTIQ 562    |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | Db          | 504 YXNDVITELASGLLGLLCKSLQKGLQRTADTEQKAVTAVFVNSKATIEDNY 593   |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | QY          | 593 RFPVPAQVQDEPQALSNMSNGVTFSL-QSLVHEVAVYTLISQAQDFSLV 651     |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | Db          | 564 KFCNEPKVREDPTVFNSTINVTPESTPLTLAGCFDTQVHKCVTHUNLIT 623     |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | QY          | 652 FSGVTFPHQVYEDTLTFPPGSGTVPMSNPLWILQCNSEPR- 700             |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | Db          | 624 IFTGSSFTLCGHEMLTLFPHQSGVTVHNVGVFWLTVGNSPNKQLAKFRDYVC 683  |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | QY          | 701 ----- 700                                                 |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | Db          | 684 INDDSDSVIITYESESSTTTTRMDGDSNENREHDEPYQLLAVSGIRFNS 743     |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | QY          | 701 -----GMITAL----- 705                                      |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | Db          | 744 LYKEDFNATLALNNSESTFPSTDRVSNSSSPGNISRAPATITPEPKLIPHEA 803  |      |                                        |                     |              |
| FT | REPEAT | 1519 | 1527 | 40. | QY          | 706 LYKESCDKNG-----DYED----- 721                              |      |                                        |                     |              |



























GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2003, 16:59:28 / Search time 53 Seconds  
(Without alignments)  
7001.353 Million cell updates/sec

Title: US-10-047-257-1

Perfect score: 7691

Sequence: 1 ATATYKLGVELSWDQKSD.....WFGQALNVEVGECAQDLT 1438

Scoring table: BLOSUM62

Gapop 10.0 / Gapext 0.5

Searched: 830525 seqs, 258032604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- 1: SP:vertebra.\*
- 2: SP:bacteria.\*
- 3: SP:fungi.\*
- 4: SP:human.\*
- 5: SP:invertebrate.\*
- 6: SP:plant.\*
- 7: SP:mico.\*
- 8: SP:organelle.\*
- 9: SP:phage.\*
- 10: SP:protozo.\*
- 11: SP:protist.\*
- 12: SP:virus.\*
- 13: SP:vertebrate.\*
- 14: SP:unclassified.\*
- 15: SP:unclassified.\*
- 16: SP:bacteriap.\*
- 17: SP:archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1          | 6272   | 81.5        | 2343   | 6     | Q18806      |
| 2          | 6231   | 81.0        | 2343   | 6     | O62730      |
| 3          | 2486.5 | 32.3        | 2119   | 13    | Q90X47      |
| 4          | 2386.5 | 31.0        | 2224   | 4     | O43737      |
| 5          | 1799   | 23.4        | 1157   | 11    | Q92024      |
| 6          | 1784   | 22.7        | 1157   | 11    | Q920H8      |
| 7          | 1784   | 22.7        | 1158   | 4     | Q90Q87      |
| 8          | 1747   | 22.7        | 1158   | 4     | Q90Q87      |
| 9          | 1710   | 22.2        | 1104   | 4     | O75130      |
| 10         | 1709.5 | 22.2        | 1084   | 11    | Q92L97      |
| 11         | 1618   | 21.0        | 1048   | 6     | O32T27      |
| 12         | 1544.5 | 20.0        | 355    | 11    | Q8B043      |
| 13         | 1544.5 | 20.0        | 355    | 11    | Q8B043      |
| 14         | 1221   | 15.9        | 782    | 4     | O75549      |
| 15         | 1221   | 15.9        | 782    | 4     | O75549      |
| 16         | 1104   | 14.4        | 216    | 4     | Q14286      |

#### ALIGNMENTS

##### RESULT 1

1186018806 PRELIMINARY; PRT; 2343 aa.  
AC Q18806;  
AC 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JUN-1998 (TrEMBLrel. 05, Last sequence update)  
DE 01-JAN-1998 (TrEMBLrel. 05, Last annotation update)  
DE Factor VIII.  
OS F8.  
GN Ctna familialis (Dog).  
GC Ctna familialis (Dog);  
OC Ctna familialis (Dog);  
OX NCBI\_TaxID=9615;  
RN [1]  
RS TRANSLATION FROM N.A.  
RA Gerson C. Morley C., Hoyle S., McGlynn L., Hough C., Kamine S.,  
RA Giles A., Lillierap D.;  
RT The canine factor VIII cDNA and 5' flanking sequence.\*;  
RT Ctna familialis (Dog) cDNA and 5' flanking sequence.\*;  
RT -1- SIMILARITY: CONTAINS 2 F8/8 TYPE C DOAINS.  
DR HSP; P04534; AAB87412.1; -;  
DR Interpro; IPR001117; Cn-oxidase.  
DR Interpro; IPR001117; Cn-oxidase.  
DR Pfam; PF00394; Cn-oxidase; 3.  
DR Pfam; PF00754; F5\_F8\_type\_C; 2.  
DR SMART; SMART031; PAS6C; 2.  
DR PROSITE; PS01286; PAS6C; 2.  
DR PROSITE; PS01286; PAS6C; 2.  
DR PROSITE; PS00079; MULTICOPPER\_OXIDASE1; 3.  
SQ SEQUENCE 2343 AA; 25529 MW; A854PA521CIB399 CRC64;

Query Match

Best Local Similarity

Matches 1245; Conservative

82; Mismatches 101; Indels 900; Gaps

4;

1 ATATYKLGVELSWDQKSD.....WFGQALNVEVGECAQDLT 1438

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||



[illegible]



```

QY 734 NNAI-----EPREFS 743
Db 798 NKVINDSTKALLETTFPONRVVALMNETDS ILIOFTVERKRSAPKPMPESEVIT 857
QY 744 QN-----PVLKCHREIRIT-----TLOD-----764
Db 858 MKFPTTEHINSLEHAIYVITETINMTWITDFSTFPDOSTORNIPTLEOTAL 917
QY 765 -----QREI-----DYD-----DTJVEWK-779
Db 918 LINGSBEPQSONSEMRJAFQELNAKODTDVDSNINYNQIFKNVPSCOTLINSKI 977
QY 780 --KEDPID-----787
Db 978 QYRSDPVLLESSTFSEMSMTWYDVSQYVKGSEKTAQSGELSTKTKYSGBILSLEI 1037
QY 788 -----787
Db 1038 POTSANLINSSEVLNINSLENSENRITLPMESNATFSDSTWTSSTATPAIDFNT 1097
QY 788 -----787
Db 1098 TFSNATFSDFNRTISQMSDSESNATLSDSNATLSDSNATLSDSNATLSDSNATLNS 1157
QY 788 -----787
Db 1158 SNATFSDSENKTFSSNATLFGVYSNTLSDFSLESBENTVILSGANDTKHSERV 1217
QY 788 -----BQNG-----782
Db 1218 SNTSQLSSESTENISLISLANSHMNDSESESEEVYLNKHSEALTLHLDQXE 1277
QY 793 -----792
Db 1278 BINGVESKELVHKELPDHKNYKYKSANSKNPKTEKEKKYQYVUKYKGYKTKYK 1337
QY 793 -----SPRSP-----797
Db 1338 SKDYKQPSHSSIFRQPSVLTGRSPVSEDELTFRVYVIGPRDNYHLYIPKQ 1397
QY 798 -----OKTRHYTIAVER 811
Db 1398 DOBAPDOLLDPHEVEYEVYDVKADVQALDTSQILLAKDQGNKTHYVILSEVEE 1457
QY 812 LMYQNSSEPHVIRNAGSVPQFQVQVPOFTQSGFTQFVRLKRNHMLGALGPVIA 871
Db 1458 ENDA-GYQORLJUKTQNTYVKNVFRILSTFISFINDKGMBOELGLGLTIA 1516
QY 872 EVDENIMVFNQASRPFYSFYSISY-----BEDROGAFKRNVPYENKTFYK 923
Db 1517 EYQVWTFVFRASRPFSLHAKNTHYKQEGSLDYDSEYVQYDQDPAVPPNTFTYH 1576
QY 924 KYQOHMAFKYCEFOKAMAFSDVLEKDVHSGLLGLPCHTNTLN--PAGRGVOTVS 981
Db 1577 TYNKSPQNSSECTYVYVNFVPSINDISLGLPILVOKGTDLCKDEPR-----E 1631
QY 982 FALFTTIDSTFVFNENRCAPQIOWEDPTFENFALNYVYDYLVA 1041
Db 1632 FULLPMPDENSLMYENRQVBERKRVH--DNPQDMLKPAING--IYSLKGLWY 1689
QY 1042 QOORVTHLNGSMENTHSIFSGVFTVYKVEYQVYALVYVYVUTFTYMLSKQI 1101
Db 1690 NTQALVHLINGSPDASVYHFOGTINELKZHRQVYVFLQGLPAILMLPSKVL 1749
QY 1102 VYVGLVCEHLNAGHSTF-----LVNSKQCPQIASHCHIDPQI 1143
Db 1750 WQSEVLSQQRQNTLTFPSEYKRYQVQVLSYTLFVAVDLPGLISGVQDEBI 1809
QY 1144 TACQVQVQKQNTLFLSQTANST--KEPPFQYKYLALPMIHTKYQDQSSLS 1202
Db 1810 TADSDRQVYFHLARNTKQANMTSSEPGQVLYHGVKNTYQVQVQVQVQV 1869
QY 1203 YISQPTIWSLODQOMTYGNSGT-----LWVFF-----QVDSISCKINFP 1248

```

```

Db 1870 PLYATLSTSTDKKMTIYKGSVDKVRANKTNKTNVFNMAQTFEGNAEYETENITFP 1929
QY 1249 PLYATLSTSTDKKMTIYKGSVDKVRANKTNKTNVFNMAQTFEGNAEYETENITFP 1307
Db 1930 PLYATLSTSTDKKMTIYKGSVDKVRANKTNKTNVFNMAQTFEGNAEYETENITFP 1989
QY 1308 TWPSKARILQJQSNARQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 1367
Db 1990 QWHEVYALJQKQANNAKQNDIQKQVLAKEVAKTIVYQVQVQVQVQVQVQVQV 2049
QY 1368 ISSQOQOQHTFQNG--KVVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 1425
Db 2050 LEYSDEBQWNTDQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 2109
QY 1426 NREVEGSE 1433
Db 2110 KILGLQD 2117

RESULT 4
IC 043737 PRELIMINARY; PRT: 2224 AA.
IC 043737;
DT 01-JUN-1998 (TRIMBrel. 06; Created)
DT 01-JUN-1998 (TRIMBrel. 06; Last sequence update)
DE FACTOR V. 043737 (TRIMBrel. 23; Last annotation update)
DE Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinoptera; Primates; Catarrhini; Hominoidea; Homo.
EN NCBI_TaxID=9606;
EN [!]-SEQUENCE FROM N.A.
RA Sasaki C. (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC [!]-SIMILARITY: CONTAINS 2 FS/8 TYPE C DOMAINS.
DR EMBL; Z95572; CAB16748.1; -.
DR HSPG; P12255; I027.
DR InterPro; IPR000432; P455.
DR InterPro; IPR000432; P455.
DR Pfam; PF00334; Cu-oxidase; 3.
DR Pfam; PF00754; F5 F8 type_C; 2.
DR PROSITE; PS01348; PASCC; 2; 1.
DR PROSITE; PS01348; PASCC; 2; 2.
DR PROSITE; PS00799; MULTICOPPER OXIDASE; 2.
SQ PROSITE 2224 AA; 251673 MW; 103110BD49609AB CRC64;

Query Match 31.04; Score 2386.51; DB 4; Length 2224;
Best Local Similarity 26.04; Pred. No. 2,2e-169;
Matches 592; Conservative 278; Mismatches 483; Indels 923; Gaps 35;

QY 3 RYVLGNVLSNDQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 61
Db 33 RQYVPAAGQVLSSTRE-----PNSLSLSVTSFKVIRVEPY-FKK 75
QY 62 AKRPMKVLQCTQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 121
Db 76 EKASTIGSLAFTVAAGVDIIVKHFKNKADPLSHFQVIRVNSLQSTLYDHFPA 135
QY 122 EKEDQVYQSGSYQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 135
Db 136 EKADVAQVIRVTHSISDSQVDPDPCVTHYTHSHENLIEDPNSLIGLPLCK 195
QY 187 GSAKQNTQ-THLK-FVADVQSGHSHTQVSDQVQVQVQVQVQVQVQVQVQV 239
Db 196 QITVGGTQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 239
QY 240 RSLPGLQCHRSYVHVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 299
Db 240 GTHDPTVCSHRIEMLLQSGSGPLSHFNQVGLQKHQVTSAILVQVSTANT 299

```



QY 300 LNUQOPLFCHLSHQDQKQVYKVSQPEBQJQKQKQBEADYDDDLTDSMDVYA 359  
 Db 300 VQPEKMIISLTPHQAQQAQYDIDKQKTKNFKITRE----- 342  
 QY 360 FDDNSPFOIQISVAKCHQVWHVITAEEDQWAFUWLPADRSYKQYIMQFQRI 419  
 Db 360 -----QREKQKWEYFJAEKVITDYAFVAMQKKAQKQHLNFSQI 387  
 QY 430 GKXKXVPMAYTDTITKREAI---QKSGILQPLXGCVGTLTILIPQKQSNRPAIT 476  
 Db 430 GKXKXKWTQVEDEST-KTKVTPNNKQDGLQI-IRAQVDFKLVKQKQKNSPIL 446  
 QY 477 PHGIT-----DYRELSRFLKQVHKLKOFFILQIBEFKXKTVTVEQPTGDPRLCT 530  
 Db 477 PHGITVSFYEVSUSFT---SGKNTMRVQVQZITFIRNIIILFDEFEMQKQCH 502  
 QY 503 RYSSPVMMEDEASGLQPLILCYKESVDQKQKQKQKQKQKQKQKQKQKQKQKQ 590  
 Db 503 RYSSVDLQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 562  
 QY 591 IQRPLPQAGQLEDFQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 649  
 Db 591 KMKCMQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 622  
 QY 650 SYFFSGTTEKHKYEDVLTFFPSGQKQKQKQKQKQKQKQKQKQKQKQKQKQ 695  
 Db 623 THFTQKSTFGKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 682  
 QY 696 -----DFKRS----- 700  
 Db 683 KQIPDDQSTVEIFPPSPSTVATKRMKRLKLEFDESDQDQYQKQKQKQKQKQ 742  
 QY 701 -----GVTL----- 705  
 Db 743 SMOKEEFNLTALQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 802  
 QY 706 -----LKVSCDKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 725  
 Db 803 ATTAGSFLRHLQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 862  
 QY 726 ----- 725  
 Db 863 AKHQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 736  
 QY 726 ----- 725  
 Db 736 KQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 736  
 QY 737 ----- 736  
 Db 983 GBSFTLANKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1042  
 QY 737 IRRPQF----- 80 744  
 Db 1043 ISRTEPLRSEAMTFSERLKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1102  
 QY 745 N----- 745  
 Db 1103 NSNDQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1162  
 QY 746 -----PPVLRH- 752  
 Db 1163 HKSFTPDLSQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1222  
 QY 753 ----- 752  
 Db 1223 SPALQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1282  
 QY 753 ----- 752  
 Db 1283 TPLSDLSQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1342  
 QY 753 -----QREITKQLSQREI----- 768

Db 1343 LQTMQSPALQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1402  
 QY 769 ----- 768  
 Db 1403 TELQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1462  
 QY 769 -----DYDDT 773  
 Db 1463 ITPYSESSQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1522  
 QY 774 ISVWKCKE-----DEPDICE-----DENQSP-----RSQKQKQKQ 804  
 Db 1523 TEIILPEEYQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1582  
 QY 805 FETIPEEYQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 859  
 Db 1583 YIAEESISQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1638  
 QY 860 RHLQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 911  
 Db 1639 BHLQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 971  
 QY 912 EHLQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1031  
 Db 1699 AVQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1758  
 QY 972 AKHQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1031  
 Db 1759 DSNNPQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1812  
 QY 1032 MDTLQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1871  
 Db 1813 Y-SJLQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1871  
 QY 1092 VPLPQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1151  
 Db 1872 LEMKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1931  
 QY 1152 WAPLQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1205  
 Db 1932 WAPLQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1265  
 QY 1206 QYIMVLSQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1291  
 Db 1992 EFPVLSQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 2051  
 QY 1266 STYELMQLKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1324  
 Db 2052 PTLRQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 2111  
 QY 1325 RVPQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1384  
 Db 2112 MOKANNQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 2171  
 QY 1385 KY-KVPOKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1438  
 Db 2172 SNVDKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 2224

## RESULT 5

DB 088783  
 ID 088783  
 AD 088783  
 DT 01-NOV-1998 (T=EMBLrel. 08, Created)  
 DT 01-NOV-1998 (T=EMBLrel. 08, Last sequence update)  
 DT 01-MAR-2003 (T=EMBLrel. 23, Last annotation update)  
 GN FS  
 OS FS  
 CC Mus musculus (Mouse)  
 CC Bukaryoca, Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Euteleostomi; Euteleostomi; Rodentia; Sciurognathi; Muridae; Mus.  
 NCBI\_FairP-101991,



|          |                                                                      |                                                          |      |    |      |                                                             |     |
|----------|----------------------------------------------------------------------|----------------------------------------------------------|------|----|------|-------------------------------------------------------------|-----|
| QY       | 1010                                                                 | NTOMEDPTKENVREINACTNIMTOLGCAQOORTNKLKLSHNSHNISHPGSHV     | 1069 | QY | 57   | HUFIKAPKSPWMLLOFTQTOARVDTVTITLONASHPVLGSHVSYNKSABQARYDD     | 116 |
| DB       | 1752                                                                 | ---RISPEKNAKNAVANGIYIN-LAGIMYCEGEVALLANAGROHIVHFGTLL     | 1808 | DB | 84   | QTYBETAEVANTCATLGFLOLNGOVITLHUKFASPVITLHIGVITYKDSBSGSL      | 143 |
| QY       | 1070                                                                 | TYRKVEKSEKALIN YQVUTVYKPKVAGI WVECL IGBHLKAGMSTFLYSNKQT  | 1129 | QY | 117  | QYOSREKEDVVPFGSHITVQVWLKENGPMASDPLCTVSYSHVDLVNLSLIGAL       | 176 |
| DB       | 1809                                                                 | LONRTKQLOQWELLGSPVITLKNASKPGWLLTGVENOVAGMOTFLIDNECON     | 1868 | DB | 144  | GSSGVKADUSVPFGSHYHMSIPSSNLTADPACILATVTHSDVANDFATLGL         | 203 |
| QY       | 1130                                                                 | PLAGASHIRFOPTLBSGQYQKAPKALHSGSYNNASTKE-----PPE-WIKUDLLA  | 1183 | QY | 177  | LVCERGESI---AAKKTQTLHRTFLITFAVDEKSHSE-----TNSLMQDSDAASR     | 226 |
| DB       | 1869                                                                 | PNGLSTGVISDLSLSEVLYTEFLRLKLNAGSNYSIEKTAIDFLIEPWQUMQX     | 1928 | DB | 204  | ITACGTLTONGSPORUDVHRTFLITFAVDEKSHSE-----TNSLMQDSDAASR       | 263 |
| QY       | 1184                                                                 | PMTHIRLITVQKQKESLSYSGPITMSLDKCKQRTKNGSTGLAVFQNVDSCKE     | 1243 | QY | 227  | AMPQKQVHSGVLEGLTCHERKSYVMHVGISHTVEHVSIFLECHITFLVNRHQS       | 286 |
| DB       | 1929                                                                 | EVVYVLTQCAKHLKSCFTFQVASSQTNQVPROSGKSVNTQNSDSTIE          | 1988 | DB | 264  | DNRHAIKGVVPPFLFELNCAKGVHAKPCKGNEIEVITAFPHQOMLSTRGHDTA       | 323 |
| QY       | 1244                                                                 | INTEPHIRIATVILHPTSHIRLIMELGCDJNSGMLGWSKASIDQITASSVFT     | 1303 | QY | 287  | EISPTFLATQTLAMDQQLFLFCHSHQDQKAVVYVDSPEEPFLKMNKONESBY        | 346 |
| DB       | 1989                                                                 | NKLDPIKSHIRLIMELGCDJNSGMLGWSKASIDQITASSVFT               | 2048 | DB | 334  | NIFPFTVPMVPMQKSGWILSCFVNSHRSQQAFTQCSMDPQV-----372           | 372 |
| QY       | 1304                                                                 | INMFAI-NSISARLHSGSNAPVNPVNPVNPVNPVNPVNPVNPVNPVNPVNPV     | 1382 | QY | 347  | DDDLTDSNDVVPFDDNSPFTQISVSANQGVPTWVHTVIAAEEDWDVAVLADPDS      | 406 |
| DB       | 2049                                                                 | SWKGVNPSLALANQVYVQVANKKNGVQVWELLKTKYATVYQCKSLSSBY        | 2108 | DB | 373  | -DQLTG-----KRVOTIQAHIEQMDYOP--LQTDART                       | 402 |
| QY       | 1363                                                                 | VKEPLASSQGHQVTFPQNGV--KVFQGHQISFTFVNSLDPFLTRKHPSQW       | 1420 | QY | 407  | YKS-----OYLNGPQRTGKYKVKVPMATVNTSTRTBALQNS--GILGRLVQ         | 455 |
| DB       | 2109                                                                 | YVLSQVQVQVNTVQKQVSSNDVNPBNSNTHKSGNSFTNPPIBSFTIRIIPFW     | 2168 | QY | 456  | RVGDTLITFLPKQASPVYHRTVDVLSLSEKRLQVH--LAKPPILOPSTFKX         | 512 |
| QY       | 1421                                                                 | HOZALRVEDEQADY 1438                                      |      | DB | 463  | RVGDTIQVVPFNRASQFSLQPHGVYKESKSTVNDGSHPKVASF-----EKVYTY      | 517 |
| DB       | 2169                                                                 | QSTALRLELFC-----DIY 2183                                 |      | QY | 513  | KVTVVECHPESGRLCTVSPNMRDLASGIGKPLCYHVESVQKQKQNMEDRM          | 572 |
| RESULT 6 |                                                                      |                                                          |      | DB | 518  | MYVPHAGFTAQFACILTWYFSADPTFTNSGLVGLFVACAGAGKQGVDFE           | 577 |
| AC       | Q92024                                                               | PRELIMINARY; PRT; 1157 AA.                               |      | DB | 573  | ULFQVPRNSHVTENKORFENAPQ-----VLEDEPE-FOASNIHSHINGVYDSE-      | 625 |
| DT       | 01-MAY-1999 (TRENBERGrel. 10, Created)                               |                                                          |      | DB | 578  | VLFTVTPDNESNTNAN-----QAQMLDSELSSEVETQDSNRWALHSGISLNP        | 631 |
| DT       | 01-MAY-1999 (TRENBERGrel. 10, Last sequence update)                  |                                                          |      | QY | 626  | QVSLCHEVAVVLTISLACTDPLSVFSGSTFRRHVVYEDTLTLPFGSGTFVNSHNP     | 685 |
| DT       | 01-MAY-1999 (TRENBERGrel. 23, Last annotation update)                |                                                          |      | DB | 632  | BLKMDQVQVHALLGLSTEDVNGVNGENTVQLOQMKKAVLPHVHTVATLQKQPD       | 691 |
| GN       | HEPHEIN.                                                             |                                                          |      | QY | 686  | GIMELCGRHSFMRGNTALLNACDQRTDYEDSYEDSAYLLSKNNATPEFSQK         | 745 |
| GN       | Mus musculus (Mouse).                                                |                                                          |      | DB | 692  | GIMELYQASREHEDQVHNVQC-----SHHQ-----722                      |     |
| OC       | NCB1_TaxID=10090.                                                    |                                                          |      | QY | 746  | PVYLQRQRITRTLDQSDQEDDITTSVEMKEDPDIYDDEQSPQSPQKTRHYF         | 805 |
| OC       | Medline-CFBI/63.                                                     |                                                          |      | DB | 723  | ---SPACHTQGRVHY 735                                         |     |
| RA       | Vulpe C.D., Kuo Y.M., Murphy T.L., Cowley L., Askwith C., Libina N., |                                                          |      | QY | 806  | TAARVRLMY-----GHSSEPVLRHBAASGVFKVQVQFDFDSTQ 851             |     |
| RA       | Huether, J., Anderson G.J.; homologous implicated in intestinal iron |                                                          |      | DB | 736  | INABELNYCPDSWELHNSSEKUSGVPLSNQGLLGSKATKAAFRHTQVTFRI 795     |     |
| RL       | Nat. Genet. 21:195-199 (1999).                                       |                                                          |      | QY | 852  | PLVCEHNEHLGSGVYAEVYVYENFQASRSTFYS-SLISYTBEDQKQAGEPK 910     |     |
| RL       | BSSEI, AF02567; XAD1605.1; +.                                        |                                                          |      | DB | 796  | FRAPGSEHGLGLRQVGLDITLVFNKASRPTYSIHAHVLESNTGQFAAE---852      |     |
| DR       | MDP; NGI:133240; Hepb.                                               |                                                          |      | QY | 911  | MYFKEKTKYPMKVOHAPATDFDKCAWATFSDVLEKSHSGILGFLVCHETNLA 970    |     |
| DR       | InterPro; IP000117; Cu-oxidase.                                      |                                                          |      | DB | 853  | ---GSLVYITMPTSPSGSPSACSVMSYVITKNTQDTSGLVHATCGNILE 908       |     |
| DR       | InterPro; IP000135; Metallo_oxidase2.                                |                                                          |      | QY | 971  | PANGHVOVEFALPTFDKTSKVTYFNNER--NCRAPAKQMEDQPTFVFNHFAING 1029 |     |
| DR       | PROSITE; PS00070; MULTICOPPER OXIDASE1, 2.                           |                                                          |      | DB | 909  | PNQGNMDRERFALLFLPDNGSYMLKENTATQVQSHSNVNLKANTYHSNKHGAL 968   |     |
| DR       | PROSITE; PS00080; MULTICOPPER OXIDASE2, 1.                           |                                                          |      | QY | 1030 | YINDTLGLVMAQOORIPVYTLASNSHNHSHSPGSHVTVYKCEVYKALVYQV 1089    |     |
| DR       | SEQUENCE 1157 AA; 129682 MW; 411BB1DA28FEC2FD CRC64;                 |                                                          |      | DB | 969  | KLVANLGLTGYTQGEVANTVLMAMQDQDTHVTHASSTLYQNGQSTADVDLFGTF 1028 |     |
| QY       | 1                                                                    | ATRYVYGVVLSHNSQNSDGLPVDAPFPVPSF-----PFTSYVYKTLVFTVY 56   |      | QY | 1090 | ETVBLPSKAGHVEVCLIGELHGHACKSLFLYSNKK 1126                    |     |
| DB       | 24                                                                   | AIRNYVLGQDMQNTAPKGRNVTNTQNTLVASSTFASGKNRIGSSYKNTVYKSD 83 |      |    |      |                                                             |     |

Query Match 22.48; Score 1789; DB 11; Length 1157;  
 Best Local Similarity 34.28; Pred. No. 9,7e-126;  
 Matches 402; Conservative 171; Mismatches 418; Indels 186; Gaps 20;

1 ATRYVYGVVLSHNSQNSDGLPVDAPFPVPSF-----PFTSYVYKTLVFTVY 56

24 AIRNYVLGQDMQNTAPKGRNVTNTQNTLVASSTFASGKNRIGSSYKNTVYKSD 83



Best Local similarity 32.2t; Pred. No. 7.8e-122; Matches 412; Conservative 191; Mismatches 451; Indels 226; Gaps 28;

QY 1 KRYVYKAVLSELYM---QSDGLGEVAVAFPPVPSF---PRYSVYKVLVIF 53  
 DB 24 ANTVYGLIYVAPKAGKGVNTWPNQDSI---VASFJASDNRRIGGTAKYIKE 80  
 QY 57 PVYHFKVIAKPRFPAKGLPTQYQVQVYVITLQNAQSHGKJAVNYSVYASGE 113  
 DB 81 YKDSYVQVAFAMGLFGLQVQVGVVILILNAPETVTHRGVYKESGSL 140  
 QY 114 YDQSHREKQVYKQVYKQVYKQVYKQVYKQVYKQVYKQVYKQVYKQVYKQV 173  
 DB 141 YPDSGSLADDSVFPQSHVYVITPESHAFDAFALVYVSHVQAFEDATGLI 200  
 QY 174 GAVLVCEGSI---AKETQTLKPTLIPVYKESHI-----GETNSMAPDAA 223  
 DB 201 GELITCKGADGSPQGGTDFLLSVDEMLNMENTATGSDAPASDE 260  
 QY 224 SANAPQMTVYVNSHGLGLGKHSVYVYVYVYVYVYVYVYVYVYVYVYVYVY 283  
 DB 261 TQESRMVATVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 320  
 QY 284 ANVLEISPTLITACTLQGLPCHLSHQHNSQVQVQVQVQVQVQVQVQVQVQV 343  
 DB 361 DVAISPTVATVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 372  
 QY 344 IDVDLDSHGVQVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 398  
 DB 373 ---DLTQ-----KURQVTEHILQVYVYVYVYVYVYVYVYVYVYVYVYV 401  
 QY 399 ---VLAQDSYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 452  
 DB 402 TQGLAEFG---SISDFKQSSSLGCTVYVYVYVYVYVYVYVYVYVYVYVYVYV 459  
 QY 453 IYGVGVTLLIFQVQASPTVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 510  
 DB 460 IYGVGVTLLIFQVQASPTVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 512  
 QY 511 ---YVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 566  
 DB 513 KKYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 572  
 QY 567 NSDQVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 619  
 DB 573 QVYKFFVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 625  
 QY 620 YVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 678  
 DB 626 FVLSNPLKNGVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 685  
 QY 679 FVLSNPLKNGVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 738  
 DB 686 IYGVGVTLLIFQVQASPTVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 718  
 QY 739 PRYSVYKVLVIFQVQVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 798  
 DB 799 KATRIYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 844  
 QY 730 QAAVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 899  
 DB 845 YDQSHREKQVYKQVYKQVYKQVYKQVYKQVYKQVYKQVYKQVYKQVYKQVYK 903  
 QY 900 YDQSHREKQVYKQVYKQVYKQVYKQVYKQVYKQVYKQVYKQVYKQVYKQVYK 949  
 DB 904 QVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 963  
 QY 850 LAME-----KGVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 902  
 DB 964 CHYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 1022

## RESULT 9

QY0058

ID QY0058

PRELIMINARY; PRT: 1158 AA.

QY0058

01-JUN-2001 (TRIMBLrel. 17; Created)

01-OCT-2002 (TRIMBLrel. 22; Last annotation update)

01-OCT-2002 (TRIMBLrel. 22; Last annotation update)

DT Rephaestin.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

Query Match 22.6t; Score 1742; DP 4; Length 1158;  
 Similarity 32.2t; Score 1742; DP 4; Length 1158;  
 Matches 412; Conservative 191; Mismatches 452; Indels 226; Gaps 28;

QY 1 KRYVYKAVLSELYM---QSDGLGEVAVAFPPVPSF---PRYSVYKVLVIF 53  
 DB 24 ANTVYGLIYVAPKAGKGVNTWPNQDSI---VASFJASDNRRIGGTAKYIKE 80  
 QY 57 PVYHFKVIAKPRFPAKGLPTQYQVQVYVITLQNAQSHGKJAVNYSVYASGE 113  
 DB 81 YKDSYVQVAFAMGLFGLQVQVGVVILILNAPETVTHRGVYKESGSL 140  
 QY 114 YDQSHREKQVYKQVYKQVYKQVYKQVYKQVYKQVYKQVYKQVYKQVYKQV 173  
 DB 141 YPDSGSLADDSVFPQSHVYVITPESHAFDAFALVYVSHVQAFEDATGLI 200  
 QY 174 GAVLVCEGSI---AKETQTLKPTLIPVYKESHI-----GETNSMAPDAA 223  
 DB 201 GELITCKGADGSPQGGTDFLLSVDEMLNMENTATGSDAPASDE 260  
 QY 224 SANAPQMTVYVNSHGLGLGKHSVYVYVYVYVYVYVYVYVYVYVYVYVYVY 283

DB 903 QKQKTLPHQSGHNSGUREFALLPFDENKSYLSENVATHSGDQPGINQDQTFLES 962  
 QY 1023 KPAVINGTMTQILQVWQDQQRWYVLSGNSNHSNHSFSGVYVYVYVYVYVYVYVY 1082  
 DB 963 KQWALNGCVAINGTMTQVQVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 1082  
 QY 1083 KQWALNGCVAINGTMTQVQVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 1142  
 DB 1023 KQWALNGCVAINGTMTQVQVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 1142  
 QY 1023 KQWALNGCVAINGTMTQVQVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 1142  
 DB 1023 KQWALNGCVAINGTMTQVQVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 1142  
 QY 1143 ITASQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 1202  
 DB 1074 YITSEKAVPRDIE---SENWVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 1202  
 QY 1203 YITSEKAVPRDIE---SENWVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 1202  
 DB 1116 SYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 1115

## RESULT 9

QY0058

ID QY0058

PRELIMINARY; PRT: 1158 AA.

QY0058

01-JUN-2001 (TRIMBLrel. 17; Created)

01-OCT-2002 (TRIMBLrel. 22; Last annotation update)

01-OCT-2002 (TRIMBLrel. 22; Last annotation update)

DT Rephaestin.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.

GN HPH.



DB 511 LGADGQGVQVEFFLFLVYDZNGHYSN-----ANQAAAMUPLRLSDEDEDFDS 563  
 QY 612 NMISINQVYDSI-QLSVCLHEVAVYVILSTGAGTQPLFLVFFSFTPKHMYEDTL 670  
 DB 564 NMMAINGFLSMLPDLCKDQGVVAMHLLGLGTETVGVHGMQVQYLOQMKRAM 623  
 QY 671 FPFSGTTPKNSNENGLATLQCHSDPNSRGNWALAVSSCDKNDGYEDYSISVLT 730  
 DB 634 FPHFVMAINGNDLGFTHYICQASHREAGMAYVNSQC-----664  
 QY 731 LSKNNALPESRQNPPLVAKHRELTTLQSLQSEEDVDYDTSVEMKEDPDIODES 790  
 DB 665 -----PQH-----668  
 QY 791 NOSTREFOKTEHYTIANLWYDNGHSS-HYVLANRAGSGV-----Pop 836  
 DB 669 -ATPRTQVQARYITMASEVENDYCPDSREWRHMRBEGKSYVITPLNKQGLSGY 727  
 QY 837 KAVYQFTQSTGCTQYVGEHLNRLGLAPPTAEVNTVNTKNSAPRYS-SL 895  
 DB 728 KAVAFREYDTGTTPRTPRQTPBEHGLTQKGEVDTLVTPKNSRPTVNAVJG 787  
 QY 896 TYIYEDQGPAPRPNVYKQVYKQVHAPATQDCEQVAVYDNDLVNDVNS 855  
 DB 788 LESTVYVLAAR-----PQVYVQNTIPRSGGPDNCAVMTYSAVDPTKMS 940  
 QY 956 GLICHLAPCHTNTLAPRCHQVTPQEPALFTPTDKSTVFTNM-BRNCARQILQMS 1014  
 DB 841 GLVQPLATCOGLTEPGRSGRNDREFFALFTPSKNSNLEEVNVAHSGQFGLNQ 900  
 QY 1015 DPTQVYFHAINGYVQTLQGLMAQOQRTWLLVSGNSNNTVSTHESCHVFTK 1074  
 DB 901 DETFLSNOCHANKSGVYVHMLTQVGVVAMVLAAGQVDYATHTFASFIYKNG 960  
 QY 1075 ERYHMAVNLVQVSTFEMLSVAGVHVEGLCHLGNKSTLFTVSNKQCTPLQMA 1134  
 DB 961 ENYEVNDVDFEGTEFVEMASNGMAMHGVTHVHAGSTLFTVFSK-----1011  
 QY 1135 SCHRPDTQASQVQAKARLALVSGSNMSTSPKNTVQVLAAPMTHINGIKQO 1194  
 DB 1012 TELSPLVITKTETKAPPHDIE--BANVQLGQIPIK--NYENLASLV 1059  
 QY 1195 AQKQFSLVSGHETVSLQKQKQVYR 1222  
 DB 1060 -----AISVTLVLAQGVVYQWR 1081  
 RESULT 11  
 ID Q9JL97  
 DC 01-OCT-2000 (TRENBelrel, 15, Created)  
 DT 01-OCT-2000 (TRENBelrel, 15, Last sequence update)  
 DT 01-DEC-2001 (TRENBelrel, 15, Last annotation update)  
 DE GPR-anchored ceruloplasmin.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 NC NCBI\_TaxId=10116;  
 SP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague Dawley; TISSUE=Brain;  
 RX MEDLINE=20127919; PubMed=10600599;  
 RT "Alternative RNA Splicing Generates a Glycophosphatidylinositol-  
 anchored Form of Ceruloplasmin in Mammalian Brain."  
 PL J. Biol. Chem. 275:4305-4310(2000).  
 DB HSP; P00450; K0C.  
 DR InterPro; IPR001117; Cu-oxidase.  
 DR InterPro; IPR002355; MulticCu\_oxidase.  
 DB Pfam; PF00394; Cu-oxidases; 3.

DR PROSITE; PS00079; MULTICOPPER\_OXIDASE1; 3.  
 DR PROSITE; PS00080; MULTICOPPER\_OXIDASE2; 1.  
 EQ SEQUENCE 1084 AA; 125745 AA; 50626AA154257C55 CQ64;  
 Query Match 22.2%; Score 1709.5; DB 11; Length 1084;  
 Best Local Similarity 32.7%; Pred. No. 4.6e-111;  
 Matches 384; Conservative 188; Mismatches 41; Indels 185; Gaps 22;  
 QY 3 KRYLIGAVLSMDQMC-SDLGEL-EVDAREPPEKPEKSEFNTSVVYKTLFVETVELFN 60  
 DB 22 KHYITGTEVMDVAGSSEGLSEVITVQSNFLANRGLRKTGKALISYDTGT 81  
 QY 61 TADKPEPMGLGLGTQLQEVYVITLQNASHPVLSHAGVSYKASGSDYDQTSQ 120  
 DB 82 KYTDMFGLGLGTVEKGVGVVHVNVASRPTFPAHGVITVYKANGVITPMYD 141  
 QY 121 REKEDVFGQSHVYVQVLEKNGPMASDPLCTTSLTSHGLVDLQSLGALLVCR 180  
 DB 142 PQADQDLPQQQLTQLVA-NKPEQSGGSHVETVYHEDVDAPASGLQVLLCK 200  
 QY 181 ESRLAKKKTQI-IRKTLVYKQSGHSHETNSLM-----QORDASNAWPPQRT 233  
 DB 201 KALHAKKENDQEVYVAVSVVDNLSMYLENNTKTCSEPTKVDMDMDPQENMTS 260  
 QY 234 VQYVNSLSPLGICHEVSYVHVYGMGTVEVSHISFTEGHFTLVNHRQASLEISFTF 293  
 DB 261 LNTQTLQMLQGLTSLSHQSHQKMEVYVKNVCSPEEQQLMNSREADYDQDLS 320  
 QY 294 LKATQTLQMLQGLTSLSHQSHQKMEVYVKNVCSPEEQQLMNSREADYDQDLS 353  
 DB 321 IDVSVAGVQVWGLSCQNLHKLQAGLQAFQDC-----NKESP--DQDQGR 368  
 QY 354 EKDVEDDDNSSEPTQISVAKKPKVWYVTAASEEDVAP-----LVLAP 402  
 DB 369 HV-----RH-----TYIAAETIIMDPAPSTPTFTTGNUTSLGS 402  
 QY 403 DHRYSYQVNLNNGPQIGKCKKVFQVATDETP--KTRERALQESGLGLPLAGVGD 459  
 DB 403 DSRVFEQ--GATLIGSTIKLVIREVTDSTFTRKQGPDEHGLIIGVIVAGVD 458  
 QY 460 TLALTPKQASPEPHYHGHIDVR-----PLYSRLPQGVHGLADPLPGSEIKYVTV 515  
 DB 459 IIVYVTHNQSQSLTQPMQVPTKPNBQTYG--PQSSSKQASHVAPKETFYEMTV 515  
 QY 516 TQVPSQSPQSLCTPLKSHQKIDALGLLCTVYESQVQKQIMSDKQVWL 575  
 DB 516 PRMGKPTADPVLCSQTSYQVGLDTFTGLGKMKCKKSLAQRQKQVDFEFLF 575  
 QY 576 VYVPSNRSKLTGLQPLNPAVQVQDPEQSNMHSINQVYSLQ--LSVCLREV 634  
 DB 576 ATVTPVNSLLELQDNRHETFAFENVQKDEDQPSNDQHSNPNFYQNLPLNCLGGS 635  
 QY 635 AVYKTLQSQDQDPLQVFSVYSKHYKVPDPLTLPFGSGTVMKNSVGLTQCHN 694  
 DB 636 IWTWTFSSAEADVHGYTFVSGTYSKGRDITANLTFHMSLTLLATPTDESGFVCLT 695  
 QY 695 SDFPRQALCALYKSCQKQDGTQEDSTENISAYLLSNNALPESRQNPPLVAKH 754  
 DB 696 THTYTGKQKQVTKQC--KQGEVDLY-----722  
 QY 755 ETRTTLQSQDEIDYDTISVNMKQEDVDYDDEKNSPESFQKNTVYATVEMLD 814  
 DB 723 -----QERTYIAAEVMD 738  
 QY 815 YKMSGSNHYVBNRAGSGV-----PQKVVVQFTQSGTQPLVQSLNKH 861  
 DB 739 YSISDNKDELHQLQVNSAFLOKSEFFQSKTKVTVTRFTSTFTRBOVDEAREH 798  
 QY 862 LGLGLPITAEVNTVQVQNSRYTFVSYLSBEDQQAQKAPKPVK--PNET 918  
 DB 799 LGLGLPITAEVNTVQVQNSRYTFVSYLSBEDQQAQKAPKPVK--PNET 918  
 QY 919 KTYVYKQVBNAPKDFQDCANAYFSDVQLEKQVHSGLGLPLVC---HYNTLAPRGR 975

DB 849 RTYIWDIPERSAGSDSPCTPMAYSTVOKYDYLGLVGLVCHSKYVYKVK--- 905  
 QY 976 QYVQSPALFFITPFTKSTFTFERNBERGCRAPCNQWEDPTKFNRPALNGYIMOTL 1035  
 DB 906 ---KMGTELLVFNBERNDONWYISDPKYNQKBEETPESKQMAKNGKMGD 943  
 QY 1036 PGLWAGQOIRYVLLKSKVSNHSHFGSHVFLVYVYKALVNLVGVGVZVEM 1095  
 DB 964 QGLNAGVQDQVYVYVWANGSIZIDWYTHFHSFGVFGKRGHRSVDVFLGVTGLMP 1023  
 QY 1096 PSKQNWREVELLGEHLHAGHSTFLVSNKOT 1129  
 DB 1024 POTGTEMLLACHVTHHAGWTVTVLPNQAS 1057

RESULT 12  
 QXMT27 PRELIMINARY; PRT; 1048 AA.  
 AC QXMT27  
 DT 01-NOV-1999 (T=EMBLrel. 12, Created)  
 DT 01-NOV-1999 (T=EMBLrel. 12, Last sequence update)  
 DT 01-NOV-1999 (T=EMBLrel. 12, Last annotation update)  
 DE Ceruloplasmin  
 OS Ovis aries (Sheep)  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Eucelostomata;  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Eucelostomata;  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Eucelostomata;  
 CC NCBI\_TaxID:9940;  
 RN 1 (1)  
 RS 1 (1)  
 RC TESTED FROM N.A.  
 EX MEDLINE=99384006; PubMed=10452945;  
 RA Lockhart P.J., Mercer J.F.B.;  
 RT Cloning and characterization of sheep ceruloplasmin cDNA.;  
 RL J Biol Chem 265:25715-25719 (1990)  
 DR EMBL; AF134814; A041477.1; .  
 DR HSP; P00450; ICM; Cu-oxidase  
 DR InterPro; IPR002355; Multicou oxidase.  
 DR Pfam; PF00394; Cu-oxidase; 3.  
 DR PROSITE; PS00079; MULTICOPPER OXIDASE1; 3.  
 DR NCBI\_Curators: MIM:180878  
 SQ SEQUENCE 1048 AA, 119125 NR, 252561760549CGB C8C64;  
 Query Match 21.0%; Score 1616; DB 6; Length 1048;  
 Best Local Similarity 37%; Fragment 1-1048; E-Value 1.0e-41;  
 Accession 378; Consensitive 177; Mismatch 417; Indels 20; Gaps 25;  
 QY 3 RYTLGAVLSWDSQDLSG---LPVQAPFPVPSFPFNTSVTKTLVFFVTLVLF 59  
 DB 22 KNYTLGTLFVNTY-ADNAKGLSVTHENSTVQENRKGISVYKAVLYTDEN 80  
 QY 60 NIAKEPFPVWGLPGLPTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTV 119  
 DB 81 RTVTEVWGLPGLPTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTV 140  
 QY 120 QREKDDVYFGGSHTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTV 179  
 DB 141 DLQKADNVQKQCLVILHNPFPQKGEESNVTFTVSHIDAPDLSGLGLIC 200  
 QY 180 REGIAKNTVTLK-FILLRVFPGHSH-----SPTNSMDQDNASARAP 229  
 DB 230 KQSLDEESKDEKPVVSNVSNVSNVSNVSNVSNVSNVSNVSNVSNVSNV 257  
 QY 230 WMTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTV 289  
 DB 258 MHSVNTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTVTV 317  
 QY 290 PTFVTLQTLQTLQTLQTLQTLQTLQTLQTLQTLQTLQTLQTLQTLQTLQTL 349  
 DB 318 PVLVLPVWAGVQGNQSNVSNVSNVSNVSNVSNVSNVSNVSNVSNVSNV 361

QY 350 LTSDSDVDFEDINSFPIQIRVAKPKPKVWVHIAEEDMDVAF-----VL 400  
 DB 362 -----SEW-----INGVNR-----YTLAEVWYAPSGIDAFTEHLR 359  
 QY 401 ADPRYSKQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 457  
 DB 400 AKCSAS---RAFFQGTTRIGSSKLYREYTSQNSQKRGPEEHLGLQVTAEV 457  
 QY 458 GDTLLIIFPQASFPNIPFPG-----TVRVLSRSLPQVGLKDPFIPQIF 509  
 DB 458 GDTVYFTHANAPLSIPGVQVDRNBECTVTSFQSPSG-SH-----VAPKGT 511  
 QY 510 KYMTVYVQGTQSPFCLRTVTSFVNNRSDLASGLPLCYKESVQVQVQVQVQV 569  
 DB 512 THRVYSVQVTVQVPLVAGNYTS---GSTQIFVTLQVAKLCEGSLANGLKND 569  
 QY 570 KXVILVSFEDNSNLTETMRPLNPAQGLQDFEFQASNTMHSITVQSLQ-LS 628  
 DB 570 RHEVLPVTFEDNSLDDNITPAPQVQENEPQESNDRKNGKVPWQKGLS 629  
 QY 629 VCUHVAWVILSGMGTFSVFFSQVTFKQVMTDTLTFSSGRTVSNMNGLM 688  
 DB 630 PCQDSVWMLFSQVGVVINGVYISNTVLSRGEBUTANLPVLSLMPQDTATP 689  
 QY 689 TLACINSDFNMTGATALLASVCDQVYVDSYEDVSYLLSQRNATERSFQNPV 748  
 DB 690 DVCVCTVDFNMTGATALLASVCDQVYVDSYEDVSYLLSQRNATERSFQNPV 748  
 QY 749 LKXHQHRTITLGSQBELDYDVTISVNMKKEPFDIDEDNSQFSQKTRVYIAA 808  
 DB 724 -----CEYTLIA 732  
 QY 809 VERLMYKNS---SEPVILNRKAGSV-----PQVYVFOFTGSGFTQVLR 855  
 DB 733 VEVMDSSNNKKEHLQEQNSNADKSEPTVGSNTVTKPDTSTFVGR 792  
 QY 856 CEHLNHLGLGPIVAREYENIYVFNQASRPTFSFSLZYEDQCAQAPRONVAP 915  
 DB 793 KXEBHLGLGLQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 845  
 QY 916 MFTKYVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 904  
 DB 846 GERTYKMLPFGHMGQSGCLVYVTVTVTVTVTVTVTVTVTVTVTVTVTV 904  
 QY 973 HGVQVTVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 1032  
 DB 905 -----IKLESLGLVDSNSWLDKNTKPKVQKAKERPNRQNLNRP 960  
 QY 1033 VTLPLVMAQQRITLWLSMNSHSHFSGHVFVTKGZYKALVLYPVQVETV 1092  
 DB 961 QNKAQVH-----HGVNVLVNSVHGVNHSFGTQVGHGTSVDFVQVTL 1009  
 QY 1093 ENLPSKAGVREVELLGEHLHAGHSTFLVSNKOT 1126  
 DB 1010 EPTFPGVLLCHVTHHAGWTVTVLPNQAS 1043

RESULT 13  
 QXMT27 PRELIMINARY; PRT; 355 AA.  
 AC QXMT27  
 DT 01-MAR-2003 (T=EMBLrel. 23, Created)  
 DT 01-MAR-2003 (T=EMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (T=EMBLrel. 23, Last annotation update)  
 DE Coagulation factor VIII (Fragment)  
 OS Mus musculus (Mouse)  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eucelostomata;  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eucelostomata;  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eucelostomata;  
 CC NCBI\_TaxID:10090;  
 RN 1 (1)  
 RS 1 (1)  
 RC TESTED FROM N.A.  
 EX MEDLINE=22354683; PubMed=12468851;





RN SEQUENCE FROM N.A.  
 RP  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBS databases.  
 RR EMBL; ALG03998; CAA19742.1; -.  
 DR HSRP; R04500; 1KCM, Cn-oxidase.  
 DR HSRP; R04500; 1KCM, Cn-oxidase.  
 DR Pfam; PF00394; Cn-oxidase; 2.  
 DR PROSITE; PS00079; MULTICOPPER\_OXIDASE1; 2.  
 RW Consatulation.  
 RY  
 SK  
 SQ

Query Match 15.48; Score 1221; Db 4; Length 782;  
 Best Local Similarity 32.78; Pred. No. 1.3e-82;  
 Matchee 292; Conservative 121; Mismatches 296; Indels 184; Gaps 18;  
 46 YKTLFETVHLNFAKPPNGLQPTQAEVDTVTTLNWSHVGHLAVGVY 105  
 Db 13 YKTYIKYKDSYDEVAQAMLGLGPVQAEVGVILHLQATPATVTHPRGVY 72  
 106 WKASGASVDQDQGRKEDVQPGQSVTVVQVGLKNGHMAQDPLCLTKYLAHVLV 165  
 Db 73 EKDSKSLVGLGSSQPLKADSDVPHGSHITNMTIPGSHATDPAFLMTYHSHVQAP 132  
 166 KQANGLCLNGLNGLGSLACKCTQTHKTLFLLFVADIKESH-----SETVNSLMQD 239  
 Db 133 RLIATGLIGLITCKG-----TDHDFLFVSDNLSHLNENIATYCSDDPASV 184  
 220 EKAAAGAPAPQVTVQVRSGLGCKHAKSVHVIQNGTTPHUSIELSEHTFLVR 279  
 Db 185 KEDETQESNEALNFGVGNLPUNLCAQVAMHLEKGNELDVHTAFPHQGLTTR 244  
 280 NHRQASLEISPTTFACATLNDLGGTLFGLCHLSKQKHQVQVQVQSPERQDLMKN 339  
 Db 245 GHRTDVANIFFATVTAENVPEECVTLISQVNSHFDQGLYKYSCKAPPV----- 300  
 340 NNEAEVDQDLDSDNGVRFDDNDSPTQIISVAKGHTWHHTIARSDHDPVAPL- 398  
 Db 301 -----DLTGT-----KVQVFIENHLOMDYGMG 325  
 399 -----VLADQRYKSYVUNGQDTCRYKCTQPMATVETKTRALQHS--QT 448  
 Db 326 HDGSTGCKLKEPC--SISDFPKSSSLGTTWVAVTAQQTQPKHLEEDHLGI 383  
 449 LQRLAYNGSTLILLNKSQASVHYVHGLTDVPLYSRLPKGVHKLQPTLKEI 508  
 Db 384 LQPIVIAEVDGTLQVNVNAGQPFPMQHVQ-----FYEDGVCTV--YNGSSYGLV 436  
 509 FK-----YKTYTVEDQPKQPCVRYVSKSFNNMERLQSLQRLAQHESVDOR 562  
 Db 437 AKPEKVTIRMTVPHHAGTADQACLTWVFSADPTIEDTNSGLVGLVCRAGALGD 496  
 563 GQVMSKSNVILISVPRSHVLFENIQTLQFIPKACVQL-----EDPS--PQASNLN 615  
 Db 497 GAGQVKGKFFLLVLEDKKSYNSR-----NQAAMLDPELLSDTEGQSDNHH 549  
 616 SINNVYDLS--QLSVTNEVAVTILSGCTDFSVFSSPTFKHGVVDTLIPFS 674  
 Db 550 ALNGFLFSLNLELDCKGSDTVAMHLLGTEIDVHGTVNQNTVQLOQNGKRAMLFET 609  
 675 GETVFNKSNFQNLGLCHNSDFNMGNTALLKVASQDNQGVYEDYEDISALLSN 734  
 Db 610 FVALNMQDNLGTFETICAGSHREAGALTNVQC----- 646  
 735 NAIPFSGQNPVFLKQKREITTTTQSQREIDYDTISVEMKEDPDIYEDBNQSP 794  
 Db 647 -----PQGL-----ATP 653  
 795 RFQCKVHTFIANVELADYQMSSS--PHYENRQAGSEV-----PQKVV 840  
 Db 654 FQRYQARYITNALEVVDYCPQSRWERHMQSEKDYGTITLSNKGGLSRYKAV 713

QY 841 KQVTCGCTQPIYQGLAHMLGLGPVIAEIVENIMATFENQASRPPVPS 893  
 DB 714 FRETDTGTFATPEKATGPKATLQPLVIGENODILVTVPFNALSHVSVMA 766

Search completed: December 9, 2003, 17:03:59

Job time : 79 secs